

Nichtparametrisches Bootstrap in heteroskedastischen Zeitreihen

Vom Fachbereich für Mathematik und Informatik
der Technischen Universität Braunschweig
genehmigte Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr.rer.nat.)

von Frank Palkowski

1. Referent: Prof. Dr. Jens-Peter Kreiß
2. Referent: Prof. Dr. Jürgen Franke

eingereicht am: 23. Juli 2004
Datum des Rigorosum: 2. Dezember 2004

Ich danke Herrn Prof. Dr. Jens-Peter Kreiß für die Anregung zu dieser Arbeit
und für die gute Betreuung.

Inhaltsverzeichnis

1	Einleitung	1
2	Schätzen in nichtparametrischen ARCH-Modellen	11
2.1	Schätzer für die bedingte Varianz	11
2.2	Stabilität	12
2.3	Asymptotische Normalität des Schätzers für die Varianz	14
2.4	Konvergenz der Ableitungen des Varianz-Schätzers	29
3	Wild Bootstrap	35
3.1	Wild Bootstrap im NARCH-Modell	35
3.2	Wild Bootstrap für die bedingte Varianz	39
4	Simulationen	52
5	Paarweises Bootstrap	61

6	Autoregressives Bootstrap	67
7	Anhang	71
7.1	Mischen	71
7.2	Martingaldifferenzschemata	73

Kapitel 1

Einleitung

Politische und soziale Unruhen, Wirtschaftskrisen, Kriege und drohende Kriege, Naturkatastrophen und Terroranschläge, so hat sich in der Vergangenheit gezeigt, wirken sich offenbar auf Finanzmärkte aus, denn in Perioden solcher Ereignisse gibt es stark schwankende Kurse an Finanzmärkten. In der Statistik bzw. der Ökonometrie sagt man dann, die Volatilität verändert sich mit der Zeit, die (unter der Vergangenheit) bedingte Varianz ist nicht konstant oder die zugrundeliegende Zeitreihe ist bedingt heteroskedastisch.

Historische Ereignisse wie z.B. der Prager Frühling 1968, die Ölkrise 1974, der Sturz des Schahs von Persien 1980, der zweite Golfkrieg 1990, die Wirtschaftskrise in Südostasien 1997, die Russische Finanzkrise 1998, die Zerstörung des World Trade Center 2001 sind für solche Schwankungen am Finanzmarkt zumindest als eine ihrer Ursachen ausgemacht wurden.

Die Ursachen des Börsencrashes am Schwarzen Montag, 19. Oktober 1987, sind bis heute nicht vollständig geklärt. Eine vielleicht etwas weit her geholte Begründung versucht die folgende Agenturmeldung von Associated Press, abgedruckt in der Seattle Times vom 8. Oktober 1987, zu geben:

WIE KAM ES ZUR RASANTEN TALFAHRT DER BÖRSENKURSE ?

New York. Der jüngste Kursverfall an der Wall Street wird unter anderem zurückgeführt auf den angeblichen Pessimismus eines achtunddreißigjährigen Yale-Absolventen im Fach Psychologie, der früher Schlagzeuger in einer Rockband war und in den achtziger Jahren zu einem der führenden Gurus in Sachen Börsenvorhersage avancierte.

Robert Prechter ist der Verfasser des *Elliott Wave Theorist*, eines eigenwilligen Bulletins für Kapitalanleger, in dem die Ansicht vertreten wird, das Börsengeschehen spiegle die Massenpsychologie wider, und Stimmungsschwankungen zwischen gut und schlecht kämen in Wellen.

Seine nach Ansicht von Börsenmaklern erstaunlich akkuraten Vorhersagen haben dem früheren Merrill-Lynch-Analysten eine beträchtliche Anhängerschaft eingebracht. Seine Prognosen basieren auf scheinbar nebensächlichen Tendenzen, die von der Länge von Frauenröcken bis zu Rocksong-Texten reichen.

Am Dienstag wurde überall gemunkelt, Prechter habe über den exklusiven Telefonservice für die 16000 Bezieher seines Mitteilungsblatts (ein Jahresabonnement beläuft sich auf 233 Dollar) kurzfristig baissierende Börsenkurse vorhergesagt.

In Börsenkreisen verlautete, diese Gerüchte hätten zu den massiven Verkäufen an der Wall Street beigetragen, die den Dow-Jones-Index um 90 Punkte absacken ließen - der größte Tagesverlust in der Geschichte dieses bekannten Aktienindexes.

Prechters Nimbus entstand in den späten siebziger Jahren, als er Aktienkäufe empfahl, unter anderem wegen eines trostlosen Songs der Punkband The Sex Pistols.

Er vertrat die Meinung, die Trübseligkeit dieses Songs signalisiere einen Tiefpunkt der öffentlichen Stimmung, dem eine baldige Verbesserung des emotionellen und wirtschaftlichen Klimas folgen müsse. Ein paar Monate später erholte sich der Markt...

In der Finanzmathematik und Finanzzeitreihenanalyse versucht man mit stochastischen und deterministischen mathematischen Modellen das zukünftige Verhalten von Finanzmärkten vorher zu sagen. Ein finanzmathematisches Modell liefert eine gute Approximation an reale Verhältnisse, falls es Vorhersagen für beobachtete Daten macht, die nur wenig von den tatsächlichen

Resultaten abweichen. Finanzdaten bestehen aus einer Zeitreihe von Preisen P_t eines Finanzderivats, die z.B. in jeder Sekunde, stündlich oder täglich aktualisiert werden. Beobachtete Preise P_t von Finanzderivaten werden als nicht stationär angenommen. Sie werden deswegen häufig in so genannte Log-returns umgeformt:

$$\begin{aligned} X_t &= \log P_t - \log P_{t-1} \\ &= \log \frac{P_t}{P_{t-1}} \\ &= \log \left(1 + \frac{P_t - P_{t-1}}{P_{t-1}} \right) \\ &\approx \frac{P_t - P_{t-1}}{P_{t-1}}, \quad t = 1, 2, \dots \end{aligned}$$

Eine Taylorentwicklung erster Ordnung ergibt also, dass die Log-returns näherungsweise identisch mit den relativen Preisänderungen sind, falls letztere klein sind. Die Log-returns, so wird allgemein angenommen, erfüllen, zumindest über nicht zu lange Zeitperioden, im Gegensatz zu den Preisen P_t die Stationaritätsannahme. Relative Preisänderungen sind nicht additiv über verschiedene Zeitperioden. Deswegen werden Log-returns, die diese Eigenschaft besitzen, diesen gegenüber häufig bevorzugt verwendet:

$$\begin{aligned} X_t + X_{t-1} &= \log P_t - \log P_{t-1} + \log P_{t-1} - \log P_{t-2} \\ &= \log P_t - \log P_{t-2} \end{aligned}$$

Die Log-returns wurden oft als eine Folge von unabhängig identisch verteilten Zufallsvariablen angenommen. Das heißt, dass der logarithmierte Preis $\log P_t$ sich gemäß einer stochastischen Irrfahrt bzw. eines Random Walk entwickelt.

Samuelson (1973) modellierte Spekulationspreise in stetiger Zeit durch eine geometrische Brownsche Bewegung. Wird dieses Modell diskretisiert, so erhält man einen Random Walk mit unabhängig identisch normalverteilten Zuwächsen für die logarithmierten Preise in diskreter Zeit. Auch Black und Scholes (1973) nahmen in ihrer Theorie für Optionspreise eine geometrische Brownsche Bewegung an:

$$P_t = P_0 \exp \left(\sigma W_t + \left(\mu - \frac{\sigma^2}{2} \right) t \right), \quad t \in [0, T],$$

wobei W_t eine Brownsche Bewegung mit Drift μ und Volatilität σ ist. Diskretisierung dieses Modells führt zu folgendem Modell der Log-returns in

diskreter Zeit:

$$\begin{aligned}X_t &= \log P_t - \log P_{t-1} \\&= \log P_0 + \sigma W_t + \left(\mu - \frac{\sigma^2}{2}\right)t - (\log P_0 + \sigma W_{t-1} + \left(\mu - \frac{\sigma^2}{2}\right)(t-1)) \\&= \mu - \frac{\sigma^2}{2} + \sigma \epsilon_t, \quad t = 1, 2, \dots,\end{aligned}$$

wobei $\epsilon_t = W_t - W_{t-1}$, also $\{\epsilon_t\}$ unabhängig identisch normalverteilt.

Die Brownsche Bewegung ist ein physikalisches Modell für die Bewegung eines kleinen Teilchens in einer Flüssigkeit und in der Physik seit Beginn des zwanzigsten Jahrhunderts untersucht worden. Albert Einstein (1905) hat einen wichtigen Beitrag zu dieser Theorie geleistet, wofür er den Nobelpreis für Physik bekam. Vor Einsteins Beitrag schlug ein junger französischer Doktorand, namens Louis Bachelier (1900), in seiner Dissertation die Brownsche Bewegung als Modell für Spekulationspreise vor.

Als man in der zweiten Hälfte des zwanzigsten Jahrhunderts Finanzdaten mit Hilfe des Computers untersuchte, wurde die Annahme einer Finanzdaten zugrundeliegenden geometrischen Brownschen Bewegung verworfen, z.B. durch Mandelbrot (1963) und Fama (1965). Ihre empirischen Studien basierten auf verschiedenen US-stock-log-return-Zeitreihen. Ihre Schlussfolgerungen waren, dass

- es eine Abhängigkeit in den Daten gibt,
- die Volatilität sich mit der Zeit verändert,
- die Verteilung der Daten heavy-tailed und asymmetrisch ist, und daher keine Normalverteilung ist.

Diese Beobachtungen sind klare Anzeichen dafür, dass ein Random Walk mit normalverteilten Zuwächsen kein sehr realistisches Modell für Log-returns von Finanzdaten ist.

Fast zwanzig Jahre nach diesen empirischen Untersuchungen stellte Engle (1982) ein Modell auf, das diese Beobachtungen berücksichtigte. Er nannte sein Modell Autoregressive Conditional Heteroscedastic (ARCH), weil die quadrierte Volatilität, d.i. die quadrierte bedingte Standardabweichung bzw.

die bedingte Varianz, nicht konstant, also heteroskedastisch, ist und die Beobachtung zum Zeitpunkt t von vergangenen Beobachtungen abhängt:

$$\begin{aligned} X_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2, \quad t \in \mathbb{Z}, \quad \alpha_0, \dots, \alpha_{p-1} \geq 0, \quad \alpha_p > 0, \end{aligned}$$

wobei $\{\epsilon_t\}$ unabhängig identisch, aber nicht notwendig normalverteilt ist. Ein Nachteil dieses Modells ist jedoch, dass es häufig nicht gut genug an die modellierten Daten angepasst ist, wenn p nicht groß ist. Außerdem ist durch dieses Modell die Form der bedingten Varianz als Polynom der quadrierten letzten p Beobachtungen $X_{t-1}^2, \dots, X_{t-p}^2$ stark einschränkend vorgegeben.

Bollerslev (1986) verallgemeinerte das ARCH-Modell mit seinem GARCH-Modell (Generalized Autoregressive Conditional Heteroscedastic):

$$\begin{aligned} X_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad t \in \mathbb{Z}, \\ \alpha_0, \dots, \alpha_{p-1} &\geq 0, \quad \alpha_p > 0, \\ \beta_1, \dots, \beta_{q-1} &\geq 0, \quad \beta_q > 0, \end{aligned}$$

wobei $\{\epsilon_t\}$ unabhängig identisch verteilt ist mit Erwartungswert null und Varianz eins. Für $q = 0$ erhält man als Spezialfall das ARCH(p) von Engle. Eine hinreichende Bedingung für Stationarität des GARCH(p, q) ist $\alpha_0 > 0$ und $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$.

Standard-ARCH und GARCH-Modelle haben den Nachteil, eine Asymmetrie der Volatilität bezüglich des Vorzeichens vergangener relativer Preisänderungen oder Log-returns nicht darstellen zu können. Das folgt einfach daraus, dass vergangene Beobachtungen nur in quadratischer Form eingehen. Negativen und positiven Preisänderungen gleichen Betrags wird dadurch derselbe Effekt auf die Volatilität zugeschrieben. Empirische Beobachtungen zeigen jedoch häufig, dass einem negativen Log-return eine größere Volatilität folgt als einem positiven Return gleichen Betrags. Diese Wirkung wird *Leverage-Effekt* genannt.

Deswegen untersuchen wir zunächst so genannte nichtparametrische ARCH-Modelle:

$$X_t = m(X_{t-1}, \dots, X_{t-p}) + \sigma(X_{t-1}, \dots, X_{t-q}) \epsilon_t, \quad t = 0, 1, 2, \dots \quad (1.1)$$

Dabei ist $\{\epsilon_t\}$ eine Folge unabhängig identisch verteilter Zufallsvariabler mit Erwartungswert null und Varianz eins. Der bedingte Erwartungswert $m : \mathbb{R}^p \rightarrow \mathbb{R}$ mit $m(x) = E[X_t | (X_{t-1}, \dots, X_{t-p}) = (x_1, \dots, x_p)]$ und die bedingte Varianz σ^2 oder Volatilität $\sigma : \mathbb{R}^q \rightarrow (0, \infty)$ mit $\sigma^2(x) = E[(X_t - E[X_t | X_{t-1}])^2 | (X_{t-1}, \dots, X_{t-q}) = (x_1, \dots, x_q)]$ sind unbekannte Funktionen. Der Einfachheit halber betrachten wir hier nur den Fall $p = q = 1$. Die nichtparametrische Modellierung hat den Vorteil, dass wir (bis auf Messbarkeit) a priori keine Annahmen über die Gestalt der bedingten Erwartung und Varianz machen. Setzt man im NARCH-Modell (1.1) $m = 0$ und $\sigma^2(X_{t-1}, \dots, X_{t-p}) = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2$, $t \in \mathbb{Z}$, $\alpha_0, \dots, \alpha_{p-1} \geq 0$, $\alpha_p > 0$, erhält man als Spezialfall das ARCH-Modell von Engle.

Als weiterer Spezialfall seien hier noch die so genannten Threshold-ARCH-Modelle (TARCH) erwähnt. In TARCH-Modellen wird die bedingte Standardabweichung (Zakoian (1994)) beziehungsweise die bedingte Varianz (Glosten et al. (1993)) durch eine stückweise lineare Funktion modelliert. Dazu unterteilt man den Definitionsbereich der bedingten Standardabweichung in disjunkte Intervalle. Um den Einfluss positiver und negativer Returns auf die Volatilität quantitativ zu unterscheiden, unterteilt man nur in zwei Intervalle, die positiven und die negativen reellen q -Tupel. Setzt man im NARCH-Modell (1.1) $m = 0$ und

$$\sigma^\delta(X_{t-1}, \dots, X_{t-q}) = \alpha_0 + \sum_{i=1}^q (\alpha_i + \alpha_i^- \mathbf{1}\{X_{t-i} < 0\}) X_{t-i}^\delta, \quad (1.2)$$

wobei $\alpha_0, \dots, \alpha_{q-1} \geq 0$, $\alpha_q > 0$, so erhält man für $\delta = 1$ und $\alpha_i^- \leq -\alpha_i$, $i = 1, \dots, q$, Zakoians und für $\delta = 2$ und $\alpha_i^- \geq -\alpha_i$, $i = 1, \dots, q$, Glostens TARCH-Modell. Anstelle einer stückweise linearen Funktion nehmen Gouriéroux und Monfort (1992) eine (stückweise konstante) Treppenfunktion als Modell für die Volatilität. Sei A_j , $j = 1, \dots, J$, eine Partition des Definitionsbereichs der bedingten Erwartung bzw. Standardabweichung. Dann erhält man das qualitative Threshold-ARCH-Modell der Ordnung eins (QTARCH(1)) aus dem NARCH(1)-Modell durch:

$$\begin{aligned} m(x) &= \sum_{j=1}^J m_j \mathbf{1}\{x \in A_j\} \\ \sigma(x) &= \sum_{j=1}^J \sigma_j \mathbf{1}\{x \in A_j\}, \end{aligned} \quad (1.3)$$

wobei m_j und $\sigma_j \geq 0$, $j = 1, \dots, J$, reelle Zahlen sind. Der bedingte Erwartungswert (1.3) wird ebenfalls als Treppenfunktion modelliert. Gouriéroux

und Monfort (1992) geben eine ausführliche Diskussion der statistischen Eigenschaften des QTARCH. Die Threshold-Modelle sind gekennzeichnet durch einen abrupten Übergang beim Überschreiten eines Schwellenwertes zwischen zwei Bereichen. Zieht man bei der Modellierung einen glatteren Übergang vor, so kann man die Indikatorfunktion in (1.2) durch eine beliebige stetige Funktion ersetzen, die für $x \rightarrow -\infty$ gegen eins und für $x \rightarrow +\infty$ gegen null konvergiert. Eine häufige Wahl ist z.B. die an der y -Achse gespiegelte logistische Funktion $F(x) = (1 + \exp(\gamma x))^{-1}$, $\gamma > 0$, $x \in \mathbb{R}$. Die Indikatorfunktion ist der Grenzwert der logistischen Funktion für $\gamma \rightarrow \infty$. Das folgende spezielle TARCH(1)-Modell wird auch asymmetrisches ARCH(1)-Modell genannt:

$$\begin{aligned}\sigma^2(X_{t-1}) &= \alpha_0 + \alpha_1((1 - \gamma_1)^2 + 4\gamma_1 \mathbf{1}\{X_{t-1} < 0\})X_{t-1}^2 \\ &= \alpha_0 + \alpha_1(1 - 2\gamma_1(1 - 2 \cdot \mathbf{1}\{X_{t-1} < 0\}) + \gamma_1^2)X_{t-1}^2 \\ &= \alpha_0 + \alpha_1(|X_{t-1}| - \gamma_1 X_{t-1})^2,\end{aligned}$$

wobei $\alpha_0 > 0$, $\alpha_1 \geq 0$ und $|\gamma_1| \leq 1$. Für $\gamma_1 = 0$ erhält man das ARCH(1)-Modell von Engle, für $\gamma_1 \neq 0$ ein Modell, das Leverage-Effekte berücksichtigt. Ist $\gamma_1 = -1$, werden nur positive Returns ($X_{t-1} > 0$) in der Volatilität σ berücksichtigt, ist $\gamma_1 = 1$, nur negative. Dieses Modell gehört zu den so genannten asymmetrischen Power-ARCH-Modellen (APARCH), die man ebenfalls aus dem NARCH-Modell (1.1) erhält durch $m = 0$ und

$$\sigma^\delta(X_{t-1}, \dots, X_{t-q}) = \alpha_0 + \sum_{i=1}^q \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta,$$

wobei $\delta, \alpha_0 > 0$, $\alpha_1, \dots, \alpha_q \geq 0$ und $|\gamma_i| \leq 1$.

Gewisse Ähnlichkeiten zum NARCH-Modell weist das folgende auf, welches in Drees und Stărică (2002), Herzel, Stărică, Tütüncü (2002), Mikosch und Stărică (2003) und Polzehl, Spokoiny, Stărică (2004) vorgestellt wurde:

$$X_{t,T} = m\left(\frac{t}{T}\right) + \sigma\left(\frac{t}{T}\right)\epsilon_{t,T}, \quad t = 0, \dots, T$$

wobei $\{\epsilon_{t,T}\}$ unabhängig identisch verteilt ist mit Erwartungswert null und Varianz eins für festes T , und die Funktionen $m : [0, 1] \rightarrow \mathbb{R}$ und $\sigma : [0, 1] \rightarrow [0, \infty)$ als stetig oder zweifach stetig differenzierbar mit beschränkten Ableitungen vorausgesetzt werden. Der Zeitparameter t wird hier auf das Einheitsintervall $[0, 1]$ normiert, um simultane Kurvenschätzung der Funktionen m und σ auf $[0, 1]$ mit Hilfe von Kernschätzern zu ermöglichen. In diesem Modell werden im Gegensatz zum NARCH-Modell die Volatilität und der Erwartungswert deterministisch modelliert. Gemeinsam ist beiden Modellen

die Heteroskedastizität, also die nicht konstante Volatilität, und deren nicht-parametrische Approximation durch einen so genannten Nadaraya-Watson-Schätzer. Im folgenden betrachten wir nur das NARCH-Modell.

Hier werden nun verschiedene Bootstrap-Verfahren vorgestellt, die die Verteilung des Nadaraya-Watson-Schätzers approximieren. Bei der klassischen Bootstrap-Methode von Efron (1979) handelt es sich um ein Stichprobenwiederholungs- bzw. Resampling-Verfahren, bei dem durch unabhängiges Ziehen (mit Zurücklegen) aus einer gegebenen Stichprobe so genannte Bootstrap-Stichproben erzeugt werden. Die mit diesen neuen Stichproben berechnete Verteilung der Bootstrap-Statistik ist ein Schätzer für die unbekannte Verteilung der ursprünglichen Statistik. Bootstrap-Verfahren können in der Praxis in Fällen angewendet werden, in denen es nur eine relativ geringe Anzahl von Beobachtungen gibt, zum Beispiel bei Messreihen, die zu teuer oder zu aufwändig sind, um häufig durchgeführt zu werden.

Im Gegensatz zum klassischen Bootstrap-Verfahren von Efron, das unabhängig identisch verteilte Zufallsvariablen als Beobachtung voraussetzt, beschäftigen wir uns in dieser Arbeit mit Bootstrap-Methoden, die auch für abhängige Zufallsvariablen konsistentes Verhalten zeigen.

In diesem Zusammenhang sei z.B. das Autoregressions- bzw. autoregressive Bootstrap aus Franke und Wendel (1992), Kreutzberger (1993) oder Franke, Kreiß, Mammen (2001) genannt. Es basiert auf der Erzeugung des Bootstrap-Prozesses

$$X_t^* = \tilde{m}(X_{t-1}^*) + \tilde{\sigma}(X_{t-1}^*)\epsilon_t^*,$$

wobei \tilde{m} und $\tilde{\sigma}$ Schätzer von m und σ sind und $\epsilon_1^*, \dots, \epsilon_T^*$ eine (unter der ursprünglichen Stichprobe X_0, \dots, X_T) bedingt unabhängig identisch verteilte Stichprobenwiederholung (Resample). Hier wird, wie man sieht, die Abhängigkeitsstruktur des ursprünglichen Prozesses vollständig nachgebildet.

Eine weitere Bootstrap-Methode für abhängige Zufallsvariablen ist das Regressions- bzw. regressive Bootstrap aus Franke, Kreiß, Mammen (2001). In diesem wird ein Regressionsmodell mit (bedingt) festem Design (X_0, \dots, X_{T-1}) aufgestellt:

$$X_t^* = \tilde{m}(X_{t-1}) + \tilde{\sigma}(X_{t-1})\epsilon_t^*,$$

wobei $\epsilon_1^*, \dots, \epsilon_T^*$ wie oben definiert ist. Auf der rechten Seite der Gleichung wird hier der ursprüngliche Prozess X_t an Stelle eines Bootstrap-Prozesses

verwendet. Es handelt sich deshalb nicht mehr wie oben um ein autoregressives, sondern um ein regressives Bootstrap-Modell. Die Abhängigkeit zwischen den Zufallsvariablen im ursprünglichen Prozess wird nicht mehr nachgebildet, damit bleibt die Struktur nur teilweise erhalten.

Im regressiven und autoregressiven Fall wurde in Franke, Kreiß, Mammen (2001) die Konsistenz der Bootstrap-Verteilung sowohl von m als auch von σ gezeigt. Ein drittes ebenfalls in der eben genannten Arbeit vorgestelltes Bootstrap-Verfahren ist das Wild Bootstrap. In diesem wird wiederum ein Regressionsmodell mit (bedingt) festem Design (X_0, \dots, X_T) aufgestellt:

$$X_t^* = \tilde{m}(X_{t-1}) + \eta_t^*,$$

wobei $\eta_1^*, \dots, \eta_T^*$ ein unabhängiges Resample mit (bedingtem) Erwartungswert null und (bedingter) Varianz $(X_t - \hat{m}_h(X_{t-1}))^2$ und \hat{m}_h der Nadaraya-Watson-Schätzer von m ist. Die Struktur des ursprünglichen Prozesses ist hier im Vergleich zu den ersten beiden Verfahren am ungenauesten nachgebildet. Für Anwendungen des Wild Bootstrap für die Konstruktion von gleichmäßigen Konfidenzbändern sei hier die Arbeit von Neumann und Kreiß (1998) erwähnt.

Das eben vorgestellte Wild-Bootstrap-Verfahren kann allerdings nur zur Schätzung der Verteilung der bedingten Erwartung m verwendet werden. Hierbei braucht σ nicht stärkere Regularitätseigenschaften zu erfüllen als etwa Stetigkeit. In ökonometrischen Anwendungen wäre es allerdings auch, wenn nicht sogar viel eher von Interesse, eine Wild-Bootstrap-Approximation für die Verteilung der bedingten Varianz σ zu haben, z.B. zum Abschätzen und Kontrollieren von Risiken, wie z.B. Kursschwankungen am Finanzmarkt.

In dieser Arbeit wird in Kapitel 3 ein solches Verfahren vorgestellt und dessen Konsistenz gezeigt. Wie schon in Franke, Kreiß, Mammen (2001) zu sehen, ist die Approximation der unbekannten Verteilung durch Wild Bootstrap zwar nicht so gut wie die durch autoregressives Bootstrap. Dafür wird hier gezeigt werden, dass das Wild Bootstrap sogar dann funktioniert, wenn dem betrachteten Prozess gar kein NARCH-Modell zugrundeliegt. Diese Voraussetzung ist allerdings für die Korrektheit des autoregressiven Bootstrap notwendig.

In Kapitel 2 wird zunächst noch einmal das Ausgangsmodell vorgestellt und nichtparametrische Nadaraya-Watson-Schätzer für die unbekannten Prozess-Funktionen m und σ präsentiert. Dann geben wir hinreichende Stabilitäts-

bedingungen für das NARCH-Modell an und zeigen die asymptotische Normalität der Verteilung des Nadaraya-Watson-Schätzers für σ . In Kapitel 2.4 wird die Konsistenz der Ableitungen des Nadaraya-Watson-Schätzers für σ nachgewiesen.

Kapitel 4 enthält einige Ergebnisse aus Simulationen von NARCH-Prozessen zur Untersuchung der Qualität von Wild-Bootstrap-Approximationen bei endlichen Stichprobengrößen.

Im Kapitel 5 wird die Konsistenz eines weiteren Bootstrap-Verfahrens für abhängige Zufallsvariablen, des so genannten paarweisen Bootstraps, für die Verteilung der bedingten Varianz σ gezeigt. Hier wird die Abhängigkeitsstruktur des ursprünglichen Prozesses insofern berücksichtigt, als hier aufeinander folgende Zufallsvariablen des Prozesses als Paare zusammengefasst werden. Die Bootstrap-Stichprobe erhält man hier durch unabhängiges Ziehen mit Zurücklegen aus der Menge der Paare.

Kapitel 6 stellt als weitere Bootstrap-Methode das oben bereits erwähnte autoregressive Bootstrap ausführlicher vor.

Kapitel 2

Schätzen in nichtparametrischen ARCH-Modellen

2.1 Schätzer für die bedingte Varianz

Wir betrachten einen bedingten heteroskedastischen autoregressiven nichtlinearen Prozess der Ordnung eins (CHARN oder NARCH(1)) der Form

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t, \quad t = 1, 2, \dots \quad (2.1)$$

Dabei sei $\{\epsilon_t\}$ eine Folge unabhängig identisch verteilter Zufallsvariabler mit Erwartungswert null und Varianz eins. Der (unbekannte) bedingte Erwartungswert $m : \mathbb{R} \rightarrow \mathbb{R}$ mit $m(x) = E[X_t | X_{t-1} = x]$ und die (unbekannte) bedingte Varianz σ^2 oder Volatilität $\sigma : \mathbb{R} \rightarrow (0, \infty)$ mit $\sigma^2(x) = E[(X_t - E[X_t | X_{t-1}])^2 | X_{t-1} = x]$ werden nichtparametrisch durch Nadaraya-Watson-Schätzer geschätzt. Für σ^2 lautet dieser wie folgt

$$\hat{\sigma}_h^2(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right) (X_t - \hat{m}_h'(X_{t-1}))^2}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right)}, \quad (2.2)$$

wobei $K : \mathbb{R} \rightarrow [0, \infty)$ mit $\int K(v)dv = 1$ Kernfunktion heißt, die so genannte Bandweite h eine positive reelle Zahl und

$$\hat{m}_h(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right) X_t}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right)} \quad (2.3)$$

ein Schätzer für m ist.
Offensichtlich gilt: $\hat{\sigma}_h^2 \geq 0$.

Der Schätzer $\hat{\sigma}_h^2$ kann auch folgendermaßen als gewichtete Summe dargestellt werden

$$\hat{\sigma}_h^2(x) = \sum_{t=1}^T w_{t-1}(x) (X_t - \hat{m}_{h'}(X_{t-1}))^2$$

wobei $w_{t-1}(x) = \frac{K\left(\frac{x-X_{t-1}}{h}\right)}{\sum_{s=1}^T K\left(\frac{x-X_{s-1}}{h}\right)}$.

Man erhält diesen lokal-konstanten Schätzer für σ dann als Lösung des gewichteten Kleinst-Quadrate-Problems (Least Squares)

$$\sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right) ((X_t - \hat{m}_{h'}(X_{t-1}))^2 - c)^2 = \min_c$$

Vorgestellt wurde dieser Schätzer in Fan und Yao (1998). Dort wurde gezeigt, dass dieser Schätzer besser ist als der Schätzer

$$\begin{aligned} \hat{\sigma}_{h,2}^2(x) &= \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right) X_t^2}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right)} - \hat{m}_h^2(x) \\ &= \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right) (X_t - \hat{m}_h(x))^2}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right)}. \end{aligned}$$

Asymptotische Normalität von $\hat{\sigma}_{h,2}^2$ wurde in Robinson (1983) gezeigt.

2.2 Stabilität

Voraussetzungen:

Der Prozess $\{X_t\}$ ist stationär und stark mischend.

Hinreichende Bedingungen hierfür sind:

(V1) Die Zeitreihe X_1, \dots, X_T ist eine Realisierung der stationären Version des NARCH-Prozesses (2.1) mit unabhängig identisch verteilten Residuen $\epsilon_1, \dots, \epsilon_T$ mit Erwartungswert null und Varianz eins. Die Residuen ϵ_t besitzen eine Dichte p_ϵ mit $\inf_{x \in C} p_\epsilon(x) > 0$ für alle kompakten Mengen C .

(V2) $\limsup_{|x| \rightarrow \infty} \frac{E|m(x) + \sigma(x)\epsilon_1|}{|x|} < 1$.

(W1) Es existiert ein $\sigma_0 > 0$, so dass $\sigma(x) \geq \sigma_0$ für alle $x \in \mathbb{R}$.

(W2) m ist in x zweimal stetig differenzierbar mit beschränkten Ableitungen.

(W3) σ ist in x zweimal stetig differenzierbar mit beschränkten Ableitungen.

(W4) p_ϵ ist in x zweimal stetig differenzierbar mit beschränkten Ableitungen und beschränkt. $E\epsilon_1^8 < \infty$.

(W5) $g, h, h' \rightarrow 0$, $Th^5 \rightarrow B^2 > 0$ für $T \rightarrow \infty$, $h' \sim T^{-\alpha'}$ mit $0 < \alpha' < \frac{1}{6}$ und $g \sim T^{-\alpha}$ mit $0 < \alpha < \frac{1}{5}$

(W6) K besitzt kompakten Träger $[-1, 1]$. K ist symmetrisch, nicht negativ und dreimal stetig differenzierbar mit $K(1) = K'(1) = 0$ und $\int K(v)dv = 1$.

Die Voraussetzungen (V1),(V2),(W1)-(W3) sind hinreichend für Stationarität und geometrische Ergodizität des Prozesses. Statt (W1)-(W3) ist dafür auch schon die Beschränktheit von m , σ und σ^{-1} auf kompakten Mengen hinreichend. Das ist eine direkte Folgerung aus den Sätzen 1 und 2 in Diebolt und Guegan (1990), vergleiche auch Meyn und Tweedie (1993) oder Doukhan (S.106/107). Wir benötigen aber für unsere Resultate diese stärkeren Voraussetzungen. Die eindeutige stationäre Verteilung des Prozesses sei π . Die Voraussetzungen stellen sicher, dass die stationäre Verteilung π der Zeitreihe $\{X_t\}$ eine positive Lebesgue-Dichte p besitzt. Aus (2.1) erhalten wir

$$p(x) = \int \frac{1}{\sigma(u)} p_\epsilon\left(\frac{x - m(u)}{\sigma(u)}\right) d\pi(u), \quad (2.4)$$

denn für B aus der Borel- σ -Algebra \mathcal{B} über \mathbb{R} gilt

$$\begin{aligned} P(x, B) &= P(X_1 \in B | X_0 = x) = P(m(x) + \sigma(x)\epsilon_1 \in B) \\ &= P(\epsilon_1 \in \{\frac{y - m(x)}{\sigma(x)} | y \in B\}) = \int_B \frac{1}{\sigma(x)} p_\epsilon(\frac{y - m(x)}{\sigma(x)}) dy \\ &\quad \text{und} \quad \pi(B) = \int P(x, B) \pi(dx) \end{aligned}$$

Aus der Stationarität und geometrischen Ergodizität des Markov-Prozesses $\{X_t\}$ aus (2.1) folgt, dass der Prozess α -mischend ist mit geometrisch fallenden Mischkoeffizienten: $\alpha(k) \leq C \cdot \rho^k$, $C > 0$, $0 < \rho < 1$ (s. Doukhan, Kapitel 2.4 und 1.3).

Die Ordnung $O(T^{-\frac{1}{5}})$ der Bandweite h ist optimal im Sinne der (kleinsten) mittleren quadratischen Abweichung. Eine zweite größere Bandweite g im Vergleich zu h in Voraussetzung (W5) für eine so genannte Überglättung des Schätzers $\hat{\sigma}_g$, so dass $Tg^5 \rightarrow \infty$, ist für eine explizite Schätzung des Erwartungswertes der im folgenden Abschnitt untersuchten Verteilung notwendig.

2.3 Asymptotische Normalität des Schätzers für die Varianz

Wir untersuchen das asymptotische Verhalten von $\mathcal{L}^\sigma(x)$, der Verteilung von $\sqrt{Th}(\hat{\sigma}_h^2(x) - \sigma^2(x))$.

Zunächst nehmen wir $m = 0$ als bekannt, d.h. $X_t = \sigma(X_{t-1})\epsilon_t$ an. Später verallgemeinern wir auf den Fall m unbekannt, wobei wir sehen werden, dass die Asymptotik unverändert bleibt. Wir spalten den zu untersuchenden Term $\sqrt{Th}(\hat{\sigma}_h^2(x) - \sigma^2(x))$ in folgende Summanden auf, den so genannten Varianz- und Erwartungswertanteil

$$\begin{aligned} & \frac{\frac{1}{\sqrt{Th}} \sum_t K(\frac{x - X_{t-1}}{h}) \sigma^2(X_{t-1}) (\epsilon_t^2 - 1)}{\frac{1}{Th} \sum_t K(\frac{x - X_{t-1}}{h})} \\ & + \frac{\frac{1}{\sqrt{Th}} \sum_t K(\frac{x - X_{t-1}}{h}) (\sigma^2(X_{t-1}) - \sigma^2(x))}{\frac{1}{Th} \sum_t K(\frac{x - X_{t-1}}{h})} \end{aligned}$$

Wir betrachten zunächst den Zähler des Varianzterms.

Lemma 2.1: Gelten (V1),(V2) und (W3)-(W6). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) \sigma^2(X_{t-1})(\epsilon_t^2 - 1) \implies \mathcal{N}(0, \tau^2(x)) ,$$

wobei $\tau^2(x) = \sigma^4(x)p(x) \int K^2(v)dv(E\epsilon_1^4 - 1)$.

Beweis: Sei $\mathcal{F}_t = \sigma(X_0, \dots, X_t) = \sigma(X_0, \epsilon_1, \dots, \epsilon_t)$ die von X_0, \dots, X_t bzw. von $X_0, \epsilon_1, \dots, \epsilon_t$ erzeugte σ -Algebra. Wir zeigen, dass die Voraussetzungen einer Version (Brown, 1971) des Zentralen Grenzwertsatzes für Martingaldifferenzschemata erfüllt werden, nämlich, dass

$$\begin{aligned} \frac{1}{Th} \sum_t E[K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1})(\epsilon_t^2 - 1)^2 | \mathcal{F}_{t-1}] \\ \rightarrow \sigma^4(x)p(x) \int K^2(v)dv(E\epsilon_1^4 - 1) \end{aligned}$$

nach Wahrscheinlichkeit, und dass für alle $\delta > 0$, wiederum nach Wahrscheinlichkeit

$$\begin{aligned} \frac{1}{Th} \sum_t E[K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1})(\epsilon_t^2 - 1)^2 \\ 1\{\frac{1}{Th} K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1})(\epsilon_t^2 - 1)^2 > \delta\} | \mathcal{F}_{t-1}] \rightarrow 0 \end{aligned}$$

Zuerst müssen wir dafür nachweisen, dass $\xi_{t+1} := \frac{1}{\sqrt{h}} K\left(\frac{x - X_t}{h}\right) \sigma^2(X_t)(\epsilon_{t+1}^2 - 1)$ ein quadratisch integrierbares Martingaldifferenzschema ist. Es gilt

$$E[\xi_{t+1} | \mathcal{F}_t] = \frac{1}{\sqrt{h}} K\left(\frac{x - X_t}{h}\right) \sigma^2(X_t) E[\epsilon_{t+1}^2 - 1 | \mathcal{F}_t] = 0 ,$$

da $\{\epsilon_t\}$ unabhängig identisch verteilt ist mit $E[\epsilon_1^2] = 1$, und

$$\begin{aligned} E\xi_{t+1}^2 &= \frac{1}{h} E K^2\left(\frac{x - X_t}{h}\right) \sigma^4(X_t) E[\epsilon_1^2 - 1]^2 \\ &= \frac{1}{h} \int K^2(v) p(x + hv) \sigma^4(x + hv) h dv E[\epsilon_1^2 - 1]^2 \\ &= \int K^2(v) p(x + hv) \sigma^4(x + hv) dv E[\epsilon_1^2 - 1]^2 < \infty \end{aligned}$$

Da K und σ in einer Umgebung von x beschränkt sind, kann die Erfüllung der bedingten Lindeberg-Bedingung mit Hilfe der Cauchy-Schwarzschen und der Cebyssevschen Ungleichung und (W4) geschlossen werden aus

$$\begin{aligned}
& \frac{1}{Th} \sum_t E[(\epsilon_t^2 - 1)^2 1\{(\epsilon_t^2 - 1)^2 > \delta Th\}] \\
&= \frac{1}{h} E[(\epsilon_t^2 - 1)^2 1\{(\epsilon_t^2 - 1)^2 > \delta Th\}] \\
&\leq \frac{1}{h} (E(\epsilon_t^2 - 1)^4)^{\frac{1}{2}} (E[1\{(\epsilon_t^2 - 1)^2 > \delta Th\}]^2)^{\frac{1}{2}} \\
&= \frac{1}{h} (E(\epsilon_t^2 - 1)^4)^{\frac{1}{2}} P((\epsilon_t^2 - 1)^2 > \delta Th)^{\frac{1}{2}} \\
&\leq \frac{1}{h} (E(\epsilon_t^2 - 1)^4)^{\frac{1}{2}} \left(\frac{1}{\delta^2 T^2 h^2} E(\epsilon_t^2 - 1)^4 \right)^{\frac{1}{2}} \\
&= O\left(\frac{1}{Th^2}\right) = o(1)
\end{aligned}$$

Nun wird die andere Bedingung überprüft. Es gilt

$$\begin{aligned}
& \frac{1}{Th} \sum_t E[K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1}) (\epsilon_t^2 - 1)^2 | \mathcal{F}_{t-1}] \\
&= \frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1}) (E\epsilon_t^4 - 1)
\end{aligned}$$

Den Erwartungswert dieser Zufallsvariablen berechnet man mit Hilfe einer Taylorentwicklung zweiter Ordnung von σ und p um x , wobei \tilde{x} eine geeignete Stelle zwischen x und X_{t-1} bezeichnet.

$$\begin{aligned}
& \frac{1}{Th} \sum_t E K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1}) (E\epsilon_1^4 - 1) \\
&= \frac{1}{h} \int K^2(v) \sigma^4(x + hv) p(x + hv) h dv (E\epsilon_1^4 - 1) \\
&= \int K^2(v) (\sigma^4(x) + hv(\sigma^4)'(x) + \frac{1}{2}(hv)^2(\sigma^4)''(\tilde{x})) \\
&\quad (p(x) + hv p'(x) + \frac{1}{2}(hv)^2 p''(\tilde{x})) dv (E\epsilon_1^4 - 1) \\
&= \sigma^4(x) p(x) (E\epsilon_1^4 - 1) \int K^2(v) dv + O(h^2)
\end{aligned}$$

denn wegen der Symmetrie von K^2 gilt: $\int v K^2(v) dv = 0$.

Diese Zufallsvariable konvergiert im quadratischen Mittel, also auch nach Wahrscheinlichkeit, gegen ihren Erwartungswert:

$$\begin{aligned}
& E\left(\frac{1}{Th} \sum_t \left[K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1})(E\epsilon_t^4 - 1) \right. \right. \\
& \quad \left. \left. - EK^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1})(E\epsilon_t^4 - 1) \right] \right)^2 \\
& \leq \frac{1}{T^2 h^2} 2 \sum_{s=1}^T \sum_{t=1}^s |Cov(K^2\left(\frac{x - X_{s-1}}{h}\right) \sigma^4(X_{s-1}), K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1}))| \\
& \quad (E\epsilon_1^4 - 1)^2 \\
& \leq \frac{const}{T^2 h^2} \sum_{s=1}^T \sum_{t=1}^s \alpha(s - t) \leq \frac{const}{T^2 h^2} \sum_{s=1}^T \sum_{l=0}^{\infty} \rho^l = O\left(\frac{1}{Th^2}\right) = o(1) \quad \square
\end{aligned}$$

Hier wurde bei der zweiten Abschätzung eine Kovarianzungleichung verwendet (siehe Anhang, Abschnitt 1).

Lemma 2.2: Gelten (V1),(V2), (W3)-(W6). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (\sigma^2(X_{t-1}) - \sigma^2(x)) \rightarrow b(x) \quad \text{n. W.},$$

wobei $b(x) = B \cdot \int v^2 K(v) dv \cdot [(\sigma^2)'(x)p'(x) + \frac{1}{2}(\sigma^2)''(x)p(x)]$.

Beweis: Eine Taylorentwicklung 2. Ordnung für die linke Seite liefert

$$\begin{aligned}
& \frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (X_{t-1} - x) (\sigma^2)'(x) \\
& + \frac{1}{2} \frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (X_{t-1} - x)^2 (\sigma^2)''(\hat{X}_{t-1})
\end{aligned}$$

wobei \hat{X}_{t-1} einen geeigneten Wert zwischen x und X_{t-1} , also, da K Träger $[-1, 1]$ besitzt, aus dem Intervall $[x - h, x + h]$, bezeichnet. Wir betrachten den ersten Summanden.

Zeige Konvergenz im quadratischen Mittel, und damit auch nach Wahrschein-

lichkeit, gegen Erwartungswert

$$\begin{aligned}
& E \left(\frac{1}{\sqrt{Th}} \sum_t \left(K\left(\frac{x - X_{t-1}}{h}\right)(X_{t-1} - x) - E\left[K\left(\frac{x - X_{t-1}}{h}\right)(X_{t-1} - x)\right] \right) \right)^2 \\
&= E \left(\frac{h}{\sqrt{Th}} \sum_t \left(K\left(\frac{x - X_{t-1}}{h}\right) \frac{X_{t-1} - x}{h} - E\left[K\left(\frac{x - X_{t-1}}{h}\right) \frac{X_{t-1} - x}{h}\right] \right) \right)^2 \\
&\leq \frac{h^2}{Th} \cdot 2 \sum_{s=1}^T \sum_{t=1}^s |Cov \left(K\left(\frac{x - X_{s-1}}{h}\right) \frac{X_{s-1} - x}{h}, K\left(\frac{x - X_{t-1}}{h}\right) \frac{X_{t-1} - x}{h} \right)| \\
&\leq \frac{const}{Th} h^2 \sum_{s=1}^T \sum_{t=1}^s \alpha(s-t) \leq \frac{h \cdot const}{T} \sum_{s=1}^T \sum_{l=0}^{\infty} \rho^l \leq h \cdot const = o(1)
\end{aligned}$$

Eine Taylorentwicklung für p und die aus der Symmetrie von K resultierende Folgerung $\int vK(v)dv = 0$ führen zu folgendem Ausdruck für den Erwartungswert

$$\begin{aligned}
& E\left[\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)(X_{t-1} - x)\right] \\
&= \sqrt{Th} EK\left(\frac{x - X_1}{h}\right) \frac{X_1 - x}{h} \\
&= \sqrt{Th^3} \int K(v)vp(x + hv)dv \\
&= \sqrt{Th^3} \int K(v)(p(x) + hvp'(x) + \frac{1}{2}(hv)^2p''(\tilde{x}))dv \\
&= \sqrt{Th^3}p(x) \int K(v)v dv + \sqrt{Th^5}p'(x) \int K(v)v^2 dv + O(\sqrt{Th^7}) \\
&\rightarrow B \cdot p'(x) \int K(v)v^2 dv
\end{aligned}$$

Analoges Vorgehen liefert unter Berücksichtigung der Stetigkeit von $(\sigma^2)''$ für den zweiten Summanden der Taylorentwicklung die Konvergenz gegen $\frac{1}{2}B \cdot \int v^2 K(v)dv \cdot p(x)(\sigma^2)''(x)$. \square

Lemma 2.3: Gelten (W4)-(W6). Dann gilt für ein $x \in \mathbb{R}$ nach Wahrscheinlichkeit:

$$\frac{1}{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) \rightarrow p(x)$$

Beweis: Berechne den Erwartungswert

$$\begin{aligned}
& E\left[\frac{1}{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)\right] \\
&= \frac{1}{h} \int K\left(\frac{x - u}{h}\right) p(u) du \\
&= \frac{1}{h} \int K(v) p(x + hv) h dv \\
&= \int K(v) (p(x) + hv p'(x) + \frac{1}{2} (hv)^2 p''(\tilde{x})) dv \\
&= p(x) \int K(v) dv + h p'(x) \int v K(v) dv + \frac{1}{2} h^2 \int v^2 K(v) p''(\tilde{x}) dv \\
&= p(x) + O(h^2)
\end{aligned}$$

Zeige Konvergenz der Zufallsvariablen gegen Erwartungswert

$$\begin{aligned}
& E\left[\frac{1}{Th} \sum_t \left(K\left(\frac{x - X_{t-1}}{h}\right) - EK\left(\frac{x - X_{t-1}}{h}\right)\right)^2\right] \\
&\leq \frac{1}{Th^2} EK^2\left(\frac{x - X_{t-1}}{h}\right) \\
&= \frac{1}{Th^2} \int K^2\left(\frac{x - u}{h}\right) p(u) du \\
&= \frac{1}{Th^2} \int K^2(v) p(x + hv) h dv = O\left(\frac{1}{Th}\right) = o(1) \quad \square
\end{aligned}$$

Erweitern wir nun auf den Fall m unbekannt, tauchen bei der Zerlegung des zu betrachtenden Termes

$$\begin{aligned}
& \sqrt{Th}(\hat{\sigma}_h^2(x) - \sigma^2(x)) \\
&= \frac{\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (X_t - m(X_{t-1}) + m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))^2}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right)} - \sigma^2(x)
\end{aligned}$$

in Summanden zusätzlich zu den beiden bisher untersuchten die folgenden auf

$$\begin{aligned}
& \frac{\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))^2}{\frac{1}{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)} \\
&+ \frac{\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) 2\sigma(X_{t-1}) \epsilon_t (m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))}{\frac{1}{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)}
\end{aligned}$$

Für den ersten gilt nach Masry/Tjostheim (1994)

$$\begin{aligned}
& \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) (m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))^2}{\sum_t K(\frac{x-X_{t-1}}{h})} \\
& \leq \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) (\sup_{u \in [x-h, x+h]} |m(u) - \hat{m}_{h'}(u)|)^2}{\sum_t K(\frac{x-X_{t-1}}{h})} \\
& = O_P(\sqrt{Th}) O_P(\frac{\log T}{Th'}) = O_P(\frac{\log T}{\sqrt{T}} T^{\alpha' - \frac{1}{10}}) = o_P(1),
\end{aligned}$$

denn $\alpha' < \frac{1}{6}$.

Für den zweiten

Lemma 2.4: Gelten (V1),(V2), (W2)-(W6). Dann gilt

$$\frac{\frac{1}{\sqrt{Th}} \sum_t K(\frac{x-X_{t-1}}{h}) 2\sigma(X_{t-1}) \epsilon_t (m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))}{\frac{1}{Th} \sum_t K(\frac{x-X_{t-1}}{h})} = o_P(1)$$

Beweis: Wir führen mit $m - \hat{m}_{h'}$ eine Taylorentwicklung 2. Ordnung um x durch. Dabei bezeichne \tilde{x} eine geeignete Stelle zwischen x und X_{t-1} . Für die weitere Abschätzung verwenden wir die Dreiecksungleichung.

$$\begin{aligned}
& \left| \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) 2\sigma(X_{t-1}) \epsilon_t (m(X_{t-1}) - \hat{m}_{h'}(X_{t-1}))}{\sum_t K(\frac{x-X_{t-1}}{h})} \right| \\
& = \left| \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) 2\sigma(X_{t-1}) \epsilon_t (m(x) - \hat{m}_{h'}(x))}{\sum_t K(\frac{x-X_{t-1}}{h})} \right. \\
& \quad + \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) 2\sigma(X_{t-1}) \epsilon_t (X_{t-1} - x) (m'(x) - \hat{m}_{h'}'(x))}{\sum_t K(\frac{x-X_{t-1}}{h})} \\
& \quad \left. + \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) 2\sigma(X_{t-1}) \epsilon_t \frac{1}{2} (X_{t-1} - x)^2 (m''(\tilde{x}) - \hat{m}_{h'}''(\tilde{x}))}{\sum_t K(\frac{x-X_{t-1}}{h})} \right|
\end{aligned}$$

$$\begin{aligned}
&\leq 2|m(x) - \hat{m}_{h'}(x)| \left| \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) \sigma(X_{t-1}) \epsilon_t}{\sum_t K(\frac{x-X_{t-1}}{h})} \right| \\
&\quad + 2|m'(x) - \hat{m}'_{h'}(x)| \left| \frac{\sqrt{Th} \sum_t K(\frac{x-X_{t-1}}{h}) (X_{t-1} - x) \sigma(X_{t-1}) \epsilon_t}{\sum_t K(\frac{x-X_{t-1}}{h})} \right| \\
&\quad + \sup_{\tilde{x} \in [x-h, x+h]} |m''(\tilde{x}) - \hat{m}''_{h'}(\tilde{x})| \\
&\quad \frac{\sqrt{Th} \sum_t |K(\frac{x-X_{t-1}}{h}) (X_{t-1} - x)^2 \sigma(X_{t-1}) \epsilon_t|}{\sum_t K(\frac{x-X_{t-1}}{h})}
\end{aligned}$$

Bekannt ist, dass $|\hat{m}_{h'}(x) - m(x)| = o_P(1)$ und dass

$$\left| \frac{1}{Th} \sum_t K(\frac{x-X_t}{h}) \sigma(X_t) \epsilon_{t+1} \right| = O_P((Th)^{-\frac{1}{2}})$$

denn $\frac{1}{\sqrt{Th}} \sum_t K(\frac{x-X_t}{h}) \sigma(X_t) \epsilon_{t+1}$ ist asymptotisch normalverteilt (siehe auch Beweis von Lemma 2.1).

Da $\frac{1}{Th} \sum_{t=1}^T K(\frac{x-X_{t-1}}{h}) \rightarrow p(x)$ nach Wahrscheinlichkeit gemäß Lemma 2.3 und $p(x) \geq C > 0$ für festes x , ist der erste Summand der Taylorentwicklung von der Ordnung $o_P(1)$, verschwindet also nach Wahrscheinlichkeit asymptotisch.

Bekannt ist aus Franke, Kreiß, Mammen (1997), dass $\hat{m}'_{h'}(x)$ nach Wahrscheinlichkeit gegen $m'(x)$ konvergiert, also zeigen wir, dass

$$\left| \frac{1}{Th} \sum_t K(\frac{x-X_t}{h}) (X_t - x) \sigma(X_t) \epsilon_{t+1} \right| = O_P(\sqrt{\frac{h}{T}})$$

Es gilt:

$$\begin{aligned}
&E \left| \frac{1}{Th} \sum_t K(\frac{x-X_t}{h}) (X_t - x) \sigma(X_t) \epsilon_{t+1} \right|^2 \\
&= \frac{1}{Th^2} E K^2(\frac{x-X_1}{h}) (X_1 - x)^2 \sigma^2(X_1) E \epsilon_2^2 \\
&= \frac{1}{Th^2} \int K^2(\frac{x-u}{h}) (u - x)^2 \sigma^2(u) p(u) du E \epsilon_2^2 \\
&= \frac{1}{Th^2} \int K^2(v) (hv)^2 \sigma^2(u) p(x + hv) h dv E \epsilon_2^2 \\
&= O(\frac{h}{T})
\end{aligned}$$

Also ist auch der zweite Summand der Taylorentwicklung von der Ordnung $o_P(1)$, verschwindet also nach Wahrscheinlichkeit asymptotisch. Um zu beweisen, dass auch der dritte Summand von der Ordnung $o_P(1)$ ist, zeigen wir zunächst

$$\frac{1}{Th} \sum_t |K(\frac{x - X_t}{h})(X_t - x)^2 \sigma(X_t) \epsilon_{t+1}| = O_P(h^2).$$

Betrachte deshalb

$$\begin{aligned} E(\frac{1}{Th} \sum_t |K(\frac{x - X_t}{h})(X_t - x)^2 \sigma(X_t) \epsilon_{t+1}|) \\ &= \frac{1}{h} E|K(\frac{x - X_1}{h})(X_1 - x)^2 \sigma(X_1)| E|\epsilon_2| \\ &= \frac{1}{h} \int K(\frac{x - u}{h})(u - x)^2 \sigma(u) p(u) du E|\epsilon_2| \\ &= \frac{1}{h} \int K(v)(hv)^2 \sigma(u) p(x + hv) h dv E|\epsilon_2| \\ &= O(h^2) \end{aligned}$$

Wir zeigen nun, dass

$$\sup_{u \in [x-h, x+h]} |\hat{m}_{h'}''(u) - m''(u)| = O_P((Th')^\beta) = o_P(1) \text{ für } \beta < 0.$$

Dazu berechnen wir zuerst die zweite Ableitung von \hat{m}'' nach x :

$$\begin{aligned} \hat{m}_{h'}''(x) &= \frac{\frac{1}{Th'^3} \sum_t K''(\frac{x - X_{t-1}}{h'}) X_t}{\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'})} \\ &\quad - 2 \frac{\frac{1}{Th'^2} \sum_t K'(\frac{x - X_{t-1}}{h'}) X_t (\frac{1}{Th'^2} \sum_t K'(\frac{x - X_{t-1}}{h'}))}{(\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'}))^2} \\ &\quad - \frac{\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'}) X_t (\frac{1}{Th'^3} \sum_t K''(\frac{x - X_{t-1}}{h'}))}{(\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'}))^2} \\ &\quad + 2 \frac{\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'}) X_t (\frac{1}{Th'^2} \sum_t K'(\frac{x - X_{t-1}}{h'}))}{(\frac{1}{Th'} \sum_t K(\frac{x - X_{t-1}}{h'}))^3} \end{aligned}$$

Das asymptotische Verhalten des Nenners ist bekannt. Für die Terme in den Zählern wird exemplarisch das asymptotische Verhalten des Terms

$$\frac{1}{Th'^3} \sum_t K''(\frac{x - X_{t-1}}{h'}) X_t$$

untersucht. Zeige:

$$\sup_{u \in [x-h, x+h]} \left| \frac{1}{Th'^3} \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - c \right| = O_P((Th')^\beta), \beta < 0, c < \infty.$$

Durch Subtrahieren und Addieren des bedingten Erwartungswertes und anschließender Anwendung der Dreiecksungleichung ergibt sich:

$$\begin{aligned} & \sup_{u \in [x-h, x+h]} \left| \frac{1}{Th'^3} \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \right| \\ & \leq \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \mid \mathcal{F}_{t-1}\right] \right| \\ & \quad + \sup_{u \in [x-h, x+h]} \left| \frac{1}{Th'^3} \sum_t E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \mid \mathcal{F}_{t-1}\right] \right| \end{aligned}$$

Für den ersten Summanden des letzten Ausdrucks lässt sich zeigen:

$$\begin{aligned} & P\left(\sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \right. \right. \\ & \quad \left. \left. - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \mid \mathcal{F}_{t-1}\right] \right| \geq (Th')^\beta \right) \rightarrow 0 \end{aligned}$$

Unterteile das Intervall $[x - h, x + h]$ dazu in äquidistante Teilintervalle I_i , $i = 1, \dots, m$, so dass $[x - h, x + h] = \bigcup_{i=1, \dots, m} [u_{i-1}, u_i]$ gilt. Die Länge der Teilintervalle ist durch m festgelegt und abhängig von T , sie wird weiter unten passend gewählt.

Somit lässt sich das Supremum über das gesamte Intervall $[x - h, x + h]$ nach oben abschätzen durch das Maximum der Suprema der einzelnen Teilintervalle $[u_{i-1}, u_i]$:

$$\begin{aligned} & P\left(\sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \mid \mathcal{F}_{t-1}\right] \right| \geq (Th')^\beta \right) \\ & \leq P\left(\max_i \sup_{u \in I_i} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t \mid \mathcal{F}_{t-1}\right] \right| \geq (Th')^\beta \right) \end{aligned}$$

$$\begin{aligned}
&\leq P(\max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t K''(\frac{u - X_{t-1}}{h'}) X_t - E[K''(\frac{u - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}] \\
&\quad - \sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - E[K''(\frac{u_i - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}]| \geq \frac{1}{2} (Th')^\beta) \\
&\quad + P(\max_i \frac{1}{Th'^3} |\sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - E[K''(\frac{u_i - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}]| \geq \frac{1}{2} (Th')^\beta) \\
&\leq P(\max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t (K''(\frac{u - X_{t-1}}{h'}) - K''(\frac{u_i - X_{t-1}}{h'})) X_t| \geq \frac{1}{4} (Th')^\beta) \\
&\quad + P(\max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t (E[K''(\frac{u_i - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}] \\
&\quad - E[K''(\frac{u - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}])| \geq \frac{1}{4} (Th')^\beta) \\
&\quad + P(\max_i \frac{1}{Th'^3} \sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - E[K''(\frac{u_i - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}]| \geq \frac{1}{2} (Th')^\beta)
\end{aligned}$$

Unter Anwendung des Mittelwertsatzes der Differentialrechnung ergibt sich mit $\tilde{u} \in [u, u_i]$ für den ersten Summanden:

$$\begin{aligned}
&\max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t (K''(\frac{u - X_{t-1}}{h'}) - K''(\frac{u_i - X_{t-1}}{h'})) X_t| \\
&\leq \max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t (K'''(\frac{\tilde{u} - X_{t-1}}{h'}) (\frac{u - X_{t-1}}{h'} - \frac{u_i - X_{t-1}}{h'})) X_t| \\
&\leq \max_i \sup_{u \in I_i} |u - u_i| \frac{1}{h'^4} \cdot \frac{1}{T} \sum_t |X_t| \cdot \sup_{\tilde{v}} |K'''(\tilde{v})| \\
&\quad (\text{Wähle die Länge der Teilintervalle so, dass gilt: } \max_i \sup_{u \in I_i} |u - u_i| \leq \frac{1}{T}.) \\
&\leq O(\frac{1}{Th'^4})
\end{aligned}$$

Damit ist der erste Summand von der Ordnung $O_P((Th')^\beta)$, $-\frac{2}{5} < \beta < 0$, denn $\frac{\frac{1}{Th'^4}}{(Th')^\beta} \rightarrow 0$ mit $h' \sim T^{-\alpha'}$, $0 < \alpha' < \frac{1}{6}$ für $-\frac{2}{5} < \beta < 0$.

Für den zweiten Summanden erhält man mit Hilfe der Rechenregeln für bedingte Erwartungswerte und wiederum des Mittelwertsatzes:

$$\begin{aligned}
&\max_i \sup_{u \in I_i} \frac{1}{Th'^3} |\sum_t (E[K''(\frac{u_i - X_{t-1}}{h'}) (m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t) | \mathcal{F}_{t-1}] \\
&\quad - E[K''(\frac{u - X_{t-1}}{h'}) (m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t) | \mathcal{F}_{t-1}])|
\end{aligned}$$

$$\begin{aligned}
&= \max_i \sup_{u \in I_i} \frac{1}{Th'^3} \left| \sum_t (K''(\frac{u_i - X_{t-1}}{h'}) - K''(\frac{u - X_{t-1}}{h'})) m(X_{t-1}) \right| \\
&\leq \max_i \sup_{u \in I_i} |u - u_i| \frac{1}{h'^4} \cdot \frac{1}{T} \sum_t |m(X_{t-1})| \cdot \sup_{\tilde{v}} |K'''(\tilde{v})| = O(\frac{1}{Th'^4})
\end{aligned}$$

Damit ist auch der zweite Summand von der Ordnung $O_P((Th')^\beta)$, $-\frac{2}{5} < \beta < 0$.

Es bleibt zu zeigen, dass auch der dritte Summand von dieser Ordnung ist. Dazu werden die Markov-, die Burkholder- (aus z.B. Hall/Heyde (1980), Theorem 2.10) und die Höldersche Ungleichung verwendet.

$$\begin{aligned}
&P(\max_i \frac{1}{Th'^3} \left| \sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - E[K''(\frac{u_i - X_{t-1}}{h'}) X_t | \mathcal{F}_{t-1}] \right| \geq \frac{1}{2} (Th')^\beta) \\
&= P(\bigcup_i \frac{1}{Th'^3} \left| \sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - E[K''(\frac{u_i - X_{t-1}}{h'}) (m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t) | \mathcal{F}_{t-1}] \right| \geq \frac{1}{2} (Th')^\beta) \\
&\leq \sum_{i=1}^m P(\frac{1}{Th'^3} \left| \sum_t K''(\frac{u_i - X_{t-1}}{h'}) X_t - K''(\frac{u_i - X_{t-1}}{h'}) m(X_{t-1}) \right| \geq \frac{1}{2} (Th')^\beta) \\
&= \sum_{i=1}^m P(\frac{1}{Th'^3} \left| \sum_t K''(\frac{u_i - X_{t-1}}{h'}) \sigma(X_{t-1}) \epsilon_t \right| \geq \frac{1}{2} (Th')^\beta) \\
&\leq \sum_{i=1}^m \frac{2^r}{(Th')^{\beta r}} (\frac{1}{Th'^3})^r E \left| \sum_t K''(\frac{u_i - X_{t-1}}{h'}) \sigma(X_{t-1}) \epsilon_t \right|^r \\
&\leq \sum_{i=1}^m \frac{2^r}{(Th')^{\beta r}} (\frac{1}{Th'^3})^r C_r E \left| \sum_t (K''(\frac{u_i - X_{t-1}}{h'}) \sigma(X_{t-1}) \epsilon_t)^2 \right|^{\frac{r}{2}} \\
&\leq C_r \frac{2^r}{(Th')^{\beta r}} (\frac{1}{Th'^3})^r T^{\frac{r}{2}} \sum_{i=1}^m E |K''^r(\frac{u_i - X_1}{h'})| \sigma^r(X_1) E |\epsilon_1|^r \\
&\leq C_r \frac{2^r}{(Th')^{\beta r}} (\frac{1}{Th'^3})^r T^{\frac{r}{2}} \cdot m \cdot h' \cdot C = O(\frac{1}{T^{\frac{r}{2}-1+\frac{1}{5}-3\alpha'r+\alpha'+\beta r-\alpha'\beta r}}) = o(1)
\end{aligned}$$

falls $(\frac{1}{2} - 3\alpha' + \beta - \alpha'\beta)r - 1 + \frac{1}{5} + \alpha' > 0$, also z.B. mit $\alpha' = \frac{1}{9}$ und $\beta = -\frac{1}{100}$ für $r \geq 5$.

Damit ist auch der dritte Summand von der Ordnung $O_P((Th')^\beta)$, $-\frac{2}{5} < \beta < 0$.

Es gilt also:

$$P\left(\sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t | \mathcal{F}_{t-1}\right] \right| \geq (Th')^\beta\right) \rightarrow 0$$

d.h.

$$\begin{aligned} \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) X_t - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) X_t | \mathcal{F}_{t-1}\right] \right| &= O_P((Th')^\beta) \\ &= o_P(1), \quad \beta < 0 \end{aligned}$$

Für den zweiten Summanden auf S. 23, Mitte erhält man

$$\begin{aligned} &\sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) (m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t) | \mathcal{F}_{t-1}\right] \right| \\ &= \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t K''\left(\frac{u - X_{t-1}}{h'}\right) m(X_{t-1}) \right| \\ &\leq \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t \left(K''\left(\frac{u - X_{t-1}}{h'}\right) m(X_{t-1}) \right. \right. \\ &\quad \left. \left. - E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) m(X_{t-1})\right] \right) \right| \\ &+ \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t E\left[K''\left(\frac{u - X_{t-1}}{h'}\right) m(X_{t-1})\right] \right| \end{aligned}$$

Nun wird mit Hilfe weiterer Zerlegungen gezeigt, dass der erste Summand des letzten Ausdrucks von der Ordnung $O_P((Th')^\beta)$ ist. Dazu wird dieselbe Intervallzerlegungsmethode angewandt wie oben und völlig analog vorgegangen. Der dritte bei dieser Zerlegung auftretende Term

$$\begin{aligned} &P\left(\max_i \frac{1}{Th'^3} \left| \sum_t \left(K''\left(\frac{u_i - X_{t-1}}{h'}\right) m(X_{t-1}) \right. \right. \right. \\ &\quad \left. \left. - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right) m(X_{t-1})\right] \right) \right| \geq (Th')^\beta\right) \\ &\leq \sum_{i=1}^m \frac{1}{(Th')^{\beta r}} \left(\frac{1}{Th'^3} \right)^r E \left| \sum_t \left(K''\left(\frac{u_i - X_{t-1}}{h'}\right) m(X_{t-1}) \right. \right. \\ &\quad \left. \left. - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right) m(X_{t-1})\right] \right) \right|^r \end{aligned}$$

lässt sich allerdings im Gegensatz zu dem entsprechenden Term bei der obigen Zerlegung nicht mit Hilfe der Burkholder-Ungleichung weiter abschätzen, da

der bedingte Erwartungswert des Ausdrucks unter der Summe ungleich null ist. Stattdessen kann hier eine Ungleichung (Theorem 2, Kapitel 1.4) aus Doukhan angewendet werden auf den Ausdruck $E|\sum_t Y_t|^r$, $r > 2$, mit der zentrierten Zufallsvariablen

$$Y_t := K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1}) - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1})\right],$$

da $Y_t \in L^{r+\delta}$ für ein $\delta > 0$ gilt, wobei L^p der Raum der messbaren Funktionen f mit $\int |f(x)|^p dx < \infty$ ist, und die dort geforderte Mischungseigenschaft aufgrund der geometrischen Ergodizität erfüllt wird.

$$\begin{aligned} E\left|\sum_t \left(K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1}) - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1})\right]\right)\right|^r \\ \leq C_r \left[\sum_t \left(E\left|K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1}) - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1})\right]\right|^{2+\delta}\right)^{\frac{r}{2+\delta}}\right]^{\frac{r}{2}} \\ = C_r \cdot T^{\frac{r}{2}} \left(E\left|K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1}) - E\left[K''\left(\frac{u_i - X_{t-1}}{h'}\right)m(X_{t-1})\right]\right|^{2+\delta}\right)^{\frac{r}{2+\delta}} \\ \leq C_r \cdot T^{\frac{r}{2}} \cdot (h')^{\frac{r}{2+\delta}} \cdot C \end{aligned}$$

Damit ergibt sich

$$\begin{aligned} \sum_{i=1}^m \frac{1}{(Th')^{\beta r}} \left(\frac{1}{Th'^3}\right)^r T^{\frac{r}{2}} \cdot (h')^{\frac{r}{2+\delta}} \cdot C_r \cdot C \\ = O\left(\frac{1}{T^{\frac{r}{2}-1+\frac{1}{5}-\frac{5+3\delta}{2+\delta}\alpha' r + \beta r - \alpha' \beta r}}\right) = o(1) \end{aligned}$$

für $(\frac{1}{2} - \frac{5+3\delta}{2+\delta}\alpha' + \beta - \alpha'\beta)r - 1 > 0$, also z.B. mit $\alpha' = \frac{1}{9}$ und $\beta = -\frac{1}{100}$ für $r \geq 5$.

Der zweite Summand ist nicht stochastisch:

$$\begin{aligned} \sup_{u \in [x-h, x+h]} \frac{1}{Th'^3} \left| \sum_t E\left[K''\left(\frac{u - X_{t-1}}{h'}\right)m(X_{t-1})\right] \right| \\ = \sup_{u \in [x-h, x+h]} \frac{1}{h'^3} |E[K''(\frac{u - X_{t-1}}{h'})m(X_{t-1})]| \\ \rightarrow |p(\hat{u})m''(\hat{u}) + 2p'(\hat{u})m'(\hat{u}) + p''(\hat{u})m(\hat{u})| = |(m \cdot p)''(\hat{u})| =: c < \infty \end{aligned}$$

für ein $\hat{u} \in [x-h, x+h]$. Denn es gilt mit den folgenden Folgerungen aus (W6)

$$\begin{aligned} \int K''(v)dv &= K'(1) - K'(-1) = 2K'(1) = 0 \\ \int vK''(v)dv &= vK'(v)|_{-1}^1 - \int K'(v)dv = -K(v)|_{-1}^1 = 0 \end{aligned}$$

$$\begin{aligned}
\int v^2 K''(v) dv &= v^2 K'(v)|_{-1}^1 - 2 \int v K'(v) dv \\
&= 0 - 2v K(v)|_{-1}^1 + 2 \int K(v) dv = 2 \\
\int v^3 K''(v) dv &= v^3 K'(v)|_{-1}^1 - 3 \int v^2 K'(v) dv \\
&= 0 - 3(v^2 K(v)|_{-1}^1 - 2 \int v K(v) dv) = 0
\end{aligned}$$

und einer Taylorentwicklung 2.Ordnung von σ^2 und p um x , wobei $\tilde{x} \in [x - h, x + h]$ sowie der Symmetrie von K'' und der Stetigkeit von m'' und p'' :

$$\begin{aligned}
&\frac{1}{h'^3} E[K''(\frac{x - X_{t-1}}{h'}) m(X_{t-1})] \\
&= \frac{1}{h'^3} \int K''(\frac{x - u}{h'}) m(u) p(u) du \\
&= \frac{1}{h'^3} \int K''(v) m(x + hv) p(x + hv) h' dv \\
&= \frac{1}{h'^2} \int K''(v) (m(x) + hv(m)'(x) + \frac{1}{2}(hv)^2(m)''(\tilde{x})) \\
&\quad (p(x) + hv p'(x) + \frac{1}{2}(hv)^2 p''(\tilde{x})) \\
&= \int (\frac{1}{2} m(x) p''(\tilde{x}) + (m)'(x) p'(x) + \frac{1}{2} (m)''(\tilde{x}) p(x)) v^2 K''(v) dv + O(h') \\
&\rightarrow (m \cdot p)''(x)
\end{aligned}$$

für ein $\tilde{x} \in [x - h, x + h]$.

Insgesamt ergibt sich damit

$$\sup_{u \in [x-h, x+h]} |\frac{1}{Th'^3} \sum_t K''(\frac{u - X_{t-1}}{h'}) X_{t-c}| = O_P((Th')^\beta), \quad -\frac{2}{5} < \beta < 0, \quad c < \infty,$$

was zu zeigen war. Analoges Vorgehen für die restlichen Terme von $\hat{m}_{h'}''(u)$ liefert entsprechende Ergebnisse. Durch Zusammenfügen dieser Teilergebnisse erhält man dann

$$\sup_{u \in [x-h, x+h]} |\hat{m}_{h'}''(u) - m''(u)| = O_P((Th')^\beta) = o_P(1), \quad -\frac{2}{5} < \beta < 0 \quad \square$$

Setzt man alle Ergebnisse aus diesem Abschnitt zusammen ergibt sich also mit Hilfe des Satzes von Slutsky

Lemma 2.5: Gelten (V1), (V2), (W2)-(W6). Für ein $x \in \mathbb{R}$ gilt dann:

$$\sqrt{Th}(\hat{\sigma}_h^2(x) - \sigma^2(x)) \implies \mathcal{N}(b(x), \tau^2(x)) \quad \text{nach Wahrscheinlichkeit,}$$

wobei $b(x) = B \cdot \int v^2 K(v) dv \cdot [(\sigma^2)'(x) \frac{p'(x)}{p(x)} + \frac{1}{2}(\sigma^2)''(x)]$

und $\tau^2(x) = \frac{\sigma^4(x)}{p(x)} \int K^2(v) dv (E\epsilon_1^4 - 1)$.

2.4 Konvergenz der Ableitungen des Varianz-Schätzers

Lemma 2.6: Gelten (V1), (V2), (W2)-(W4) und (W6), und sei $g \sim T^{-\alpha}$, $0 < \alpha < \frac{1}{6}$. Für ein $x \in \mathbb{R}$ gilt dann:

$$\begin{aligned} (i) \quad & (\hat{\sigma}_g^2)'(x) \rightarrow (\sigma^2)'(x) \\ (ii) \quad & (\hat{\sigma}_g^2)''(x) \rightarrow (\sigma^2)''(x) \\ (iii) \quad & \sup_{u \in [x-h, x+h]} |(\hat{\sigma}_g^2)''(u) - (\sigma^2)''(u)| \rightarrow 0 \\ (iv) \quad & \sup_{u \in [x-h, x+h]} |\hat{\sigma}_g^2(u) - \sigma^2(u)| \rightarrow 0 \end{aligned}$$

nach Wahrscheinlichkeit.

Beweis: Nach ein- bzw. zweimaligem Ableiten des Schätzers $\hat{\sigma}_g^2$ erhalten wir:

$$\begin{aligned} (\hat{\sigma}_g^2)'(x) &= \frac{\frac{1}{Tg^2} \sum_{t=1}^T K'(\frac{x-X_{t-1}}{g})(X_t - \hat{m}_g(X_{t-1}))^2}{\frac{1}{Tg} \sum_{t=1}^T K(\frac{x-X_{t-1}}{g})} \\ &\quad - \frac{\frac{1}{Tg} \sum_{t=1}^T K(\frac{x-X_{t-1}}{g})(X_t - \hat{m}_g(X_{t-1}))^2 (\frac{1}{Tg^2} \sum_{t=1}^T K'(\frac{x-X_{t-1}}{g}))}{(\frac{1}{Tg} \sum_{t=1}^T K(\frac{x-X_{t-1}}{g}))^2} \end{aligned}$$

und

$$\begin{aligned}
(\hat{\sigma}_g^2)''(x) &= \frac{\frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x-X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2}{\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)} \\
&- 2 \frac{\frac{1}{Tg^2} \sum_{t=1}^T K'\left(\frac{x-X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2 \left(\frac{1}{Tg^2} \sum_{t=1}^T K'\left(\frac{x-X_{t-1}}{g}\right)\right)}{\left(\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)\right)^2} \\
&- \frac{\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2 \left(\frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x-X_{t-1}}{g}\right)\right)}{\left(\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)\right)^2} \\
&+ 2 \frac{\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2 \left(\frac{1}{Tg^2} \sum_{t=1}^T K'\left(\frac{x-X_{t-1}}{g}\right)\right)}{\left(\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{g}\right)\right)^3}
\end{aligned}$$

Bekannt ist

$$\frac{1}{Tg} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{g}\right) \rightarrow p(x) \text{ nach Wahrscheinlichkeit für } T \rightarrow \infty.$$

Analog kann man zeigen

$$\begin{aligned}
\frac{1}{Tg^2} \sum_{t=1}^T K'\left(\frac{x - X_{t-1}}{g}\right) &\rightarrow p'(x) \text{ nach Wahrscheinlichkeit für } T \rightarrow \infty, \\
\frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) &\rightarrow p''(x) \text{ nach Wahrscheinlichkeit für } T \rightarrow \infty.
\end{aligned}$$

Wir zeigen nun nur noch, dass

$$\frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2 \rightarrow (\sigma^2 p)''(x) \text{ n. W. für } T \rightarrow \infty.$$

Analog lässt sich

$$\frac{1}{Tg^2} \sum_{t=1}^T K'\left(\frac{x - X_{t-1}}{g}\right)(X_t - \hat{m}_g(X_{t-1}))^2 \rightarrow (\sigma^2 p)'(x) \text{ n. W. für } T \rightarrow \infty$$

zeigen.

$$\begin{aligned}
& \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) (X_t - m(X_{t-1}) + m(X_{t-1}) - \hat{m}_g(X_{t-1}))^2 \\
&= \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) (X_t - m(X_{t-1}))^2 \\
&\quad + \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) (m(X_{t-1}) - \hat{m}_g(X_{t-1}))^2 \\
&\quad + \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) 2\sigma(X_{t-1})\epsilon_t(m(X_{t-1}) - \hat{m}_g(X_{t-1})) \\
&= \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) \sigma^2(X_{t-1})\epsilon_t^2 + o_p(1)
\end{aligned}$$

Das asymptotische Verschwinden des dritten Summanden wird analog wie Lemma 2.4 gezeigt. Für den zweiten gilt wiederum mit Masry, Tjostheim (1994):

$$\begin{aligned}
& \left| \frac{1}{Tg^3} \sum_{t=1}^T K''\left(\frac{x - X_{t-1}}{g}\right) (m(X_{t-1}) - \hat{m}_g(X_{t-1}))^2 \right| \\
&\leq \frac{1}{Tg^3} \sum_{t=1}^T |K''\left(\frac{x - X_{t-1}}{g}\right)| \left(\sup_{u \in [x-h, x+h]} |m(u) - \hat{m}_g(u)| \right)^2 \\
&= O_P\left(\frac{\ln T}{Tg^3}\right) = o_p(1)
\end{aligned}$$

Die folgenden schon im Beweis von Lemma 2.4 verwendeten Folgerungen aus (W6)

$$\begin{aligned}
\int K''(v) dv &= 0 \\
\int v K''(v) dv &= 0 \\
\int v^2 K''(v) dv &= 2 \\
\int v^3 K''(v) dv &= 0
\end{aligned}$$

und die Taylorentwicklung 2.Ordnung von σ^2 und p um x , wobei $\tilde{x} \in [x - g, x + g]$ sowie die Symmetrie von K'' und die Stetigkeit von σ'' und p'' benutzen wir bei der nachfolgenden Umformung.

Berechne den bedingten Erwartungswert

$$\begin{aligned}
& \frac{1}{Tg^3} \sum_{t=2}^{T+1} E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}] \\
&= \frac{1}{Tg^3} \sum_{t=2}^{T+1} E[K''(\frac{x - m(X_{t-2}) + \sigma(X_{t-2})\epsilon_{t-1}}{g}) \\
&\quad \sigma^2(m(X_{t-2}) + \sigma(X_{t-2})\epsilon_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}] \\
&= \frac{1}{Tg^3} \sum_{t=2}^{T+1} \int K''(v)\sigma^2(x + gv)p_\epsilon(\frac{x - m(X_{t-2})}{\sigma(X_{t-2})} + \frac{gv}{\sigma(X_{t-2})})\frac{1}{\sigma(X_{t-2})}g dv \\
&\quad E[\epsilon_t^2 | \mathcal{F}_{t-2}] \\
&= \frac{1}{Tg^2} \sum_{t=2}^{T+1} \int K''(v)(\sigma^2(x) + gv(\sigma^2)'(x) + \frac{1}{2}(gv)^2(\sigma^2)''(\tilde{x})) \\
&\quad (p_\epsilon(\frac{x - m(X_{t-2})}{\sigma(X_{t-2})}) + gvp'_\epsilon(\frac{x - m(X_{t-2})}{\sigma(X_{t-2})}) + \frac{1}{2}(gv)^2p''_\epsilon(\frac{\tilde{x} - m(X_{t-2})}{\sigma(X_{t-2})}))dv \\
&= [\frac{1}{2}\sigma^2(x)\frac{1}{T} \sum_{t=2}^{T+1} p''_\epsilon(\frac{\tilde{x} - m(X_{t-2})}{\sigma(X_{t-2})}) + (\sigma^2)'(x)\frac{1}{T} \sum_{t=2}^{T+1} p'_\epsilon(\frac{x - m(X_{t-2})}{\sigma(X_{t-2})}) \\
&\quad + \frac{1}{2}(\sigma^2)''(\tilde{x})\frac{1}{T} \sum_{t=2}^{T+1} p_\epsilon(\frac{x - m(X_{t-2})}{\sigma(X_{t-2})})] \int v^2 K''(v)dv + O(g^2) \\
&\rightarrow (\frac{1}{2}\sigma^2(x)p''(\tilde{x}) + (\sigma^2)'(x)p'(x) + \frac{1}{2}(\sigma^2)''(\tilde{x})p(x)) \int v^2 K''(v)dv \\
&= (\sigma^2 p)''(x)
\end{aligned}$$

Zeige Konvergenz gegen den bedingten Erwartungswert

$$E \left(\frac{1}{Tg^3} \sum_{t=1}^T \left(K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 - E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}] \right) \right)^2$$

$$\begin{aligned}
&\leq \frac{1}{T^2 g^6} \sum_{t=1}^T E \left(K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 - E[K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 | \mathcal{F}_{t-2}] \right)^2 \\
&\quad + \frac{1}{T^2 g^6} \sum_{t=1}^T \sum_{s=1}^{t-1} E[(K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 \\
&\quad - E[K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 | \mathcal{F}_{s-2}]) \\
&\quad (K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 - E[K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 | \mathcal{F}_{t-2}])] \\
&\leq \frac{1}{T^2 g^6} \sum_{t=1}^T E \left(K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 \right)^2 \\
&= \frac{1}{T g^6} E \left(K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \right)^2 E \epsilon_t^4 \\
&= \frac{1}{T g^6} \int (K'')^2(v) \sigma^4(x + gv) p(x + gv) g dv E \epsilon_t^4 \\
&= O\left(\frac{1}{T g^5}\right)
\end{aligned}$$

denn

$$\begin{aligned}
&E[(K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 - E[K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 | \mathcal{F}_{s-2}]) \\
&(K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 - E[K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 | \mathcal{F}_{t-2}])] \\
&= E[K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 \\
&\quad (K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 - E[K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 | \mathcal{F}_{t-2}]) \\
&\quad - E[K'' \left(\frac{x - X_{s-1}}{g} \right) \sigma^2(X_{s-1}) \epsilon_s^2 | \mathcal{F}_{s-2}]) \\
&\quad (K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 - E[K'' \left(\frac{x - X_{t-1}}{g} \right) \sigma^2(X_{t-1}) \epsilon_t^2 | \mathcal{F}_{t-2}])]
\end{aligned}$$

$$\begin{aligned}
&= E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 \\
&\quad - E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}]] \\
&\quad - E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 | \mathcal{F}_{s-2}]] \\
&\quad - E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 | \mathcal{F}_{s-2}] E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}]] \\
&= 0 - E[E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 | \mathcal{F}_{s-2}]] \\
&\quad - E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 | \mathcal{F}_{s-2}]] \\
&\quad E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}] | \mathcal{F}_{s-2}]] \\
&= E[E[K''(\frac{x - X_{s-1}}{g})\sigma^2(X_{s-1})\epsilon_s^2 | \mathcal{F}_{s-2}]] \\
&\quad (E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{s-2}]] \\
&\quad - E[E[K''(\frac{x - X_{t-1}}{g})\sigma^2(X_{t-1})\epsilon_t^2 | \mathcal{F}_{t-2}] | \mathcal{F}_{s-2}])]) \\
&= 0
\end{aligned}$$

An dieser Stelle sieht man, dass eine Überglättung des Schätzers $\hat{\sigma}_g$ im Sinne, dass $Tg^5 \rightarrow \infty$, notwendig ist, damit auch die zweite Ableitung dieses Schätzers noch schwach konsistent ist.

Teil (iii) und (iv) des Satzes beweist man analog wie $\sup_{u \in [x-h, x+h]} |\hat{m}_h''(u) - m''(u)| = O_P((Th)^\beta)$ im vorigen Kapitel. \square

Kapitel 3

Wild Bootstrap

3.1 Wild Bootstrap im NARCH-Modell

Wir stellen nun eine Resampling-Methode zur Schätzung der Verteilung von $\sqrt{Th}(\hat{\sigma}_h^2(x) - \sigma^2(x))$ vor, das so genannte Wild Bootstrap oder Münchhausenverfahren. Dieses Verfahren basiert auf der Schätzung der Verteilung der Residuen. Beim Wild Bootstrap werden T verschiedene Verteilungen durch nur T Beobachtungen geschätzt. Dieses Verfahren wurde von Wu (1986) eingeführt, für Erweiterungen und Anwendungen siehe z.B. Mammen (1992). In Härdle und Mammen (1993) wird das Wild Bootstrap in nichtparametrischer Regression benutzt, für weitere Anwendungen siehe auch Franke, Kreiß und Mammen (1996).

Der quadrierte (bedingt) zentrierte Prozess $\{(X_t - m(X_{t-1}))^2\}$ kann in folgender Weise ebenfalls als NARCH(1)-Prozess dargestellt werden. Die prozessdefinierende Gleichung (2.1) lässt sich auch schreiben als:

$$(X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}) = \sigma^2(X_{t-1})(\epsilon_t^2 - 1) =: \eta_t.$$

Es gilt offenbar $E\eta_t = 0$, $Var\eta_t = E[\sigma^2(X_{t-1})(\epsilon_t^2 - 1)]^2 = E\sigma^4(X_{t-1})(E\epsilon_t^4 - 1)$. Die Residuen $\{\eta_t\}$ sind unkorreliert, aber nicht unabhängig identisch verteilt. Außerdem gilt mit $\mathcal{F}_{t-1} = \sigma(X_0, \epsilon_1, \dots, \epsilon_{t-1})$ und wegen der aus der stochastischen Unabhängigkeit der $\{\epsilon_t\}$ resultierenden Unabhängigkeit von ϵ_t von \mathcal{F}_{t-1} und wegen $E\epsilon_t^2 = 1$

$$E[\eta_t | \mathcal{F}_{t-1}] = E[\sigma^2(X_{t-1})(\epsilon_t^2 - 1) | \mathcal{F}_{t-1}] = \sigma^2(X_{t-1})E[\epsilon_t^2 - 1 | \mathcal{F}_{t-1}] = 0$$

Daher handelt es sich bei $\{\eta_t\}$ um ein Martingaldifferenzschema.

Beim Wild Bootstrap geht man nun folgendermaßen weiter vor: Berechne die empirischen Residuen $\hat{\eta}_t := (X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1})$. Erzeuge ein unabhängiges Resample $\{\eta_t^*\}$, wobei für den (bedingten) Erwartungswert $E^*\eta_t^* = 0$, für die Varianz $E^*\eta_t^{*2} = \hat{\eta}_t^2$ und für das vierte Moment $E^*\eta_t^{*4} = \hat{\eta}_t^4$, $t = 1, \dots, T$ gelte. Dabei bezeichne E^* den bedingten Erwartungswert, gegeben X_0, \dots, X_T . Dieses ist beispielsweise zu erreichen durch Erzeugung einer unabhängig identisch verteilten Stichprobe (Sample) $\bar{\eta}_1, \dots, \bar{\eta}_T$ mit Erwartungswert null und Varianz eins und Definition von $\eta_t^* = \hat{\eta}_t \bar{\eta}_t$. Zusätzlich werde ein endliches viertes Moment von $\bar{\eta}_t$ vorausgesetzt. Der Name Wild Bootstrap kommt daher, dass die erzeugte Stichprobe $\bar{\eta}_1, \dots, \bar{\eta}_T$ von der ursprünglichen Stichprobe X_1, \dots, X_T unabhängig sein kann.

Die den NARCH-Prozess definierende Gleichung kann man mit Hilfe des anfangs dieses Kapitels definierten Residuums $\eta_t = \sigma^2(X_{t-1})(\epsilon_t^2 - 1)$ auch schreiben als:

$$(X_t - m(X_{t-1}))^2 = \sigma^2(X_{t-1}) + \eta_t.$$

Die Wild Bootstrap-Stichprobe (Resample) wird in Analogie zur prozessdefinierenden Gleichung wie folgt definiert

$$X_t^{2*} = \hat{\sigma}_g^2(X_{t-1}) + \eta_t^*. \quad (3.1)$$

Wir bilden in dieser Wild-Bootstrap-Methode also nicht die ursprüngliche Stichprobe X_1, \dots, X_T nach, sondern die Quadrate der zentrierten Zufallsvariablen. Dadurch wird erreicht, dass die (geschätzte) Regressionsfunktion, also die (geschätzte) bedingte Varianz mit der (unter X_0, \dots, X_T bedingten) Erwartung der Wild-Bootstrap-Beobachtung X_t^{2*} übereinstimmt:

$$E^* X_t^{2*} = E^*(\hat{\sigma}_g^2(X_{t-1}) + \eta_t^*) = E^* \hat{\sigma}_g^2(X_{t-1}) + E^* \eta_t^* = \hat{\sigma}_g^2(X_{t-1})$$

Den Bootstrap-Schätzer für die Varianz σ^2 definieren wir nun folgendermaßen

$$\hat{\sigma}_h^{*2}(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right) X_t^{2*}}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right)}. \quad (3.2)$$

Offenbar kann dieser Schätzer im Gegensatz zum Nadaraya-Watson-Schätzer auch negative Werte annehmen, da auch die in den Schätzer eingehende Wild-Bootstrap-Stichprobe $\{X_t^{2*}\}$ aufgrund ihrer Konstruktion im Gegensatz zu X_t^2 negative Werte annehmen kann.

Die bedingte Verteilung von $\sqrt{Th}\{\hat{\sigma}_h^{*2}(x) - \hat{\sigma}_g^2(x)\}$, gegeben X_0, \dots, X_T , wird mit $\mathcal{L}_{WB}^\sigma(x)$ bezeichnet. Dieses ist der Bootstrap-Schätzer von $\mathcal{L}^\sigma(x)$, der Verteilung von $\sqrt{Th}\{\hat{\sigma}_h^2(x) - \sigma^2(x)\}$.

Wir kommen nun zum Hauptergebnis dieses Kapitels. Wir zeigen die Konsistenz des Bootstrap-Schätzers $\mathcal{L}_{WB}^\sigma(x)$.

Satz 3.1: Gelten (V1), (V2), (W1) - (W6). Dann gilt für ein $x \in \mathbb{R}$:

$$d_K(\mathcal{L}_{WB}^\sigma(x), \mathcal{L}^\sigma(x)) \rightarrow 0 \quad \text{nach Wahrscheinlichkeit}$$

d_K bezeichnet hier den Kolmogorov-Abstand, d.h. für zwei Verteilungen P und Q ist der Abstand $d_K(P, Q)$ definiert als $\sup_{x \in \mathbb{R}} |P(X \leq x) - Q(X \leq x)|$

Beweis: Wir spalten den zu untersuchenden Term $\sqrt{Th}(\hat{\sigma}_h^{*2}(x) - \hat{\sigma}_g^2(x))$ wie im vorigen Abschnitt in einen Varianz- und Erwartungswertanteil auf

$$\frac{\sqrt{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)(X_t^{2*} - \hat{\sigma}_g^2(X_{t-1}))}{\sum_t K\left(\frac{x - X_{t-1}}{h}\right)} + \frac{\sqrt{Th} \sum_t K\left(\frac{x - X_{t-1}}{h}\right)(\hat{\sigma}_g^2(X_{t-1}) - \hat{\sigma}_g^2(x))}{\sum_t K\left(\frac{x - X_{t-1}}{h}\right)}$$

Mit den beiden folgenden Lemmata und den Ergebnissen aus dem vorigen Abschnitt ist der Satz bewiesen. Gemäß 3.1 gilt: $X_t^{2*} - \hat{\sigma}_g^2(X_{t-1}) = \eta_t^*$. Für den Varianzterm gilt:

Lemma 3.2: Gelten (V1),(V2) und (W3)-(W6). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* \implies \mathcal{N}(0, \tau^2(x)) \quad \text{nach Wahrscheinlichkeit,}$$

wobei $\tau^2(x) = \sigma^4(x)p(x) \int K^2(v)dv(E\epsilon_1^4 - 1)$.

Beweis: Wir benutzen zum Beweis den Zentralen Grenzwertsatz für Dreiecksschemata von unabhängigen Beobachtungen.

$$\begin{aligned} & E^*\left(\frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^*\right) \\ &= \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}}{h}\right) ((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1})) E \bar{\eta}_t = 0 \end{aligned}$$

Für die Varianz erhalten wir

$$\begin{aligned}
& E^* \left(\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* \right)^2 \\
&= \frac{1}{Th} \left(\sum_t E^* \left[K^2\left(\frac{x - X_{t-1}}{h}\right) (\eta_t^*)^2 \right] \right. \\
&\quad \left. + \sum_{t \neq s} E^* \left[K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* K\left(\frac{x - X_{s-1}}{h}\right) \eta_s^* \right] \right) \\
&= \frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1}) \right)^2 E \bar{\eta}_t^2 \\
&= \frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left((X_t - m(X_{t-1}) + o_P(1))^2 - \sigma^2(X_{t-1}) + o_P(1) \right)^2 \\
&= \frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left(\sigma^2(X_{t-1}) (\epsilon_t^2 - 1) \right)^2 + o_P(1) \\
&\rightarrow \sigma^4(x) p(x) \int K^2(v) dv (E \epsilon_1^4 - 1) \quad \text{nach Wahrscheinlichkeit,}
\end{aligned}$$

denn

$$\begin{aligned}
& E \left[\frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left(\sigma^2(X_{t-1}) (\epsilon_t^2 - 1) \right)^2 \right] \\
&= \frac{1}{Th} \sum_t E K^2\left(\frac{x - X_{t-1}}{h}\right) \sigma^4(X_{t-1}) E (\epsilon_t^4 - 1) \\
&= \sigma^4(x) p(x) \int K^2(v) dv (E \epsilon_1^4 - 1) + O(h^2)
\end{aligned}$$

und

$$\begin{aligned}
& E \left[\frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left(\sigma^2(X_{t-1}) (\epsilon_t^2 - 1) \right)^2 \right. \\
&\quad \left. - E \left[\frac{1}{Th} \sum_t K^2\left(\frac{x - X_{t-1}}{h}\right) \left(\sigma^2(X_{t-1}) (\epsilon_t^2 - 1) \right)^2 \right] \right] = o(1)
\end{aligned}$$

gemäß Beweis von Lemma 2.1.

Die Lindeberg-Bedingung ist erfüllt

$$\begin{aligned}
& \sum_t E^* \left[\left(\frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* \right)^2 1\left\{ \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* > \delta \right\} \right] \\
& \leq \left(E^* \left(\frac{1}{\sqrt{h}} K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* \right)^4 E^* 1\left\{ \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}}{h}\right) \eta_t^* > \delta \right\} \right)^{\frac{1}{2}} \\
& = \frac{1}{h} \left(K^4\left(\frac{x - X_{t-1}}{h}\right) \left((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1}) \right)^4 \right. \\
& \quad \left. E \bar{\eta}_t^4 P^* \left(K^2\left(\frac{x - X_{t-1}}{h}\right) \eta_t^{*2} > \delta Th \right) \right)^{\frac{1}{2}} \\
& \leq \frac{1}{h} \left(K^4\left(\frac{x - X_{t-1}}{h}\right) \left((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1}) \right)^4 \right. \\
& \quad \left. E \bar{\eta}_t^4 \frac{1}{\delta^2 T^2 h^2} E^* K^4\left(\frac{x - X_{t-1}}{h}\right) \eta_t^{*4} \right)^{\frac{1}{2}} \\
& = O_P\left(\frac{1}{Th^2}\right) = o_P(1)
\end{aligned}$$

Lemma 3.3: Gelten (V1),(V2) und (W3)-(W6). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (\hat{\sigma}_g^2(X_{t-1}) - \hat{\sigma}_g^2(x)) \rightarrow b(x) \quad (\text{nach Wahrscheinlichkeit}),$$

wobei $b(x) = B \cdot \int v^2 K(v) dv \cdot [(\sigma^2)'(x)p'(x) + \frac{1}{2}(\sigma^2)''(x)p(x)]$.

Beweis: Da $(\hat{\sigma}_g^2)'$ und $(\hat{\sigma}_g^2)''$ nach Lemma 2.6 (i) und (iii) konsistente Schätzer von $(\sigma^2)'$ und $(\sigma^2)''$ sind, verläuft der Beweis analog zu dem von Lemma 2.2.

3.2 Wild Bootstrap für die bedingte Varianz

In diesem Kapitel soll gezeigt werden, dass $\mathcal{L}^\sigma(x)$, die Verteilung von $\sqrt{Th} \{\hat{\sigma}_h^2(x) - \sigma^2(x)\}$ auch ohne die Modellannahme (1.1) konsistent durch Wild Bootstrap geschätzt werden kann.

Zur Abkürzung setzen wir $\epsilon_t := \frac{X_t - m(X_{t-1})}{\sigma(X_{t-1})}$. Dabei sind m und σ wie in Kapitel 1 als $m(x) = E[X_t | X_{t-1} = x]$ und $\sigma^2(x) = E[(X_t - m(X_{t-1}))^2 | X_{t-1} = x]$

$x]$ definiert. Die ϵ_t sind nun nicht mehr wie vorher unabhängig, allerdings gilt wie in Kapitel 2.2:

$$\begin{aligned} E\epsilon_t &= E[E[\frac{X_t - m(X_{t-1})}{\sigma(X_{t-1})} | X_{t-1}]] \\ &= E[\frac{1}{\sigma(X_{t-1})}(E[X_t | X_{t-1}] - m(X_{t-1}))] = 0 \\ Var\epsilon_t &= E[\frac{1}{\sigma^2(X_{t-1})}E[(X_t - m(X_{t-1}))^2 | X_{t-1}]] = 1 \end{aligned}$$

Die Annahmen, die wir jetzt treffen, sind die Voraussetzungen

(W1) Es existiert ein $\sigma_0 > 0$, so dass $\sigma(x) \geq \sigma_0$.

(W2) m ist zweimal stetig differenzierbar in x mit beschränkten Ableitungen.

(W3) σ ist zweimal stetig differenzierbar in x mit beschränkten Ableitungen.

(W5) $g, h \rightarrow 0$, $Th^5 \rightarrow B^2 > 0$ und $g \sim T^{-\alpha}$ mit $0 < \alpha < \frac{1}{5}$ für $T \rightarrow \infty$

(W6) K besitzt kompakten Träger $[-1, 1]$. K ist symmetrisch, nicht negativ und dreimal stetig differenzierbar mit $K(1) = K'(1) = 0$ und $\int K(v)dv = 1$

wie im modellabhängigen Fall. Außerdem wird vorausgesetzt, dass der Prozess $\{X_t\}$ stark mischend mit geometrisch fallenden Mischkoeffizienten und stationär mit (stationärer) in x strikt positiver, zweimal stetig differenzierbarer Dichte p mit beschränkten Ableitungen ist und

(W7) Das bedingte vierte Moment von ϵ_t $\kappa(x) := E[\epsilon_t^4 | X_{t-1} = x]$ ist stetig in x .

Satz 3.4: Gelten (W1) - (W7). Dann gilt für ein $x \in \mathbb{R}$:

$$d_K(\mathcal{L}_{WB}^\sigma(x), \mathcal{L}^\sigma(x)) \rightarrow 0 \quad \text{nach Wahrscheinlichkeit}$$

Beweis: Zunächst gehen wir genauso vor wie im modellabhängigen Fall, d.h. wir zerlegen den zu untersuchenden Term $\sqrt{Th}\{\hat{\sigma}_h^2(x) - \sigma^2(x)\}$ in Varianz- und Biasanteil. Für den Varianzanteil gilt:

Lemma 3.5: Gelten (W1)-(W7). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) \sigma^2(X_{t-1})(\epsilon_t^2 - 1) \implies \mathcal{N}(0, \tau^2(x)) ,$$

wobei $\tau^2(x) = \sigma^4(x)p(x) \int K^2(v)dv(E[\epsilon_1^4|X_0 = x] - 1)$.

Zum Beweis der asymptotischen Normalität des Varianzterms benutzen wir unter diesen (schwächeren) Voraussetzungen einen zentralen Grenzwertsatz für stark mischende Zufallsvariablen aus Politis/Romano/Wolf (1999): Theorem B.0.1 im Spezialfall $a = 1$, denn den zentralen Grenzwertsatz für Martingaldifferenzschemata können wir hier nicht wie im Fall eines zugrundeliegenden NARCH(1)-Prozesses benutzen, da $\eta_t = \sigma^2(X_{t-1})(\epsilon_t^2 - 1)$ hier kein Martingaldifferenzschema mehr ist.

Sei $\|X\|_p = (E|X|^p)^{\frac{1}{p}}$.

Theorem B.0.1 (Romano/Politis/Wolf, 1999): Sei $\{X_{T,t} : 1 \leq t \leq d_T\}$ ein Dreiecksschema zentrierter Zufallsvariablen mit zugehörigen Mischkoeffizienten α_T . Gelten für ein $\delta > 0$:

(i) $\|X_{T,t}\|_{2+2\delta} \leq \Delta$ für alle T, t ,

(ii) $\sigma_{T,k_T}^2 := \text{Var} \frac{1}{\sqrt{k_T}} \sum_{t=1}^{k_T} X_{T,t} \rightarrow \tilde{\sigma}^2 > 0$ für alle Folgen $\{k_T\}$, die mit $T \rightarrow \infty$ gegen ∞ konvergieren,

(iii) $\sum_{k=0}^{\infty} (k+1)^2 \alpha_T^{\frac{\delta}{4+\delta}}(k) \leq K$,

wobei Δ und K von T und k_T unabhängige Konstanten sind,

so gilt:

$$\frac{1}{\sqrt{d_T}} \sum_{t=1}^{d_T} X_{T,t} \implies \mathcal{N}(0, \tilde{\sigma}^2).$$

Das Dreiecksschema lautet in unserem Fall:

$$\begin{aligned} X_{T,t} &:= \frac{1}{\sqrt{h}} K\left(\frac{X_{t-1} - x}{h}\right) \sigma^2(X_{t-1})(\epsilon_t^2 - 1) \\ &= \frac{1}{\sqrt{h}} K\left(\frac{X_{t-1} - x}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1})), \quad 1 \leq t \leq T. \end{aligned}$$

Dieses Dreiecksschema erfüllt allerdings Bedingung (i) nicht. Es lässt sich lediglich die Erfüllung der schwächeren Bedingung

$$(i)' \quad \|X_{T,t}\|_{2+2\delta} \leq \Delta T^{\frac{1}{5} \frac{\delta}{2+\delta}} \text{ für alle } T, t$$

zeigen, denn:

$$\begin{aligned} & \|X_{T,t}\|_{2+2\delta} \\ &= (E|\frac{1}{\sqrt{h}}K(\frac{X_{t-1}-x}{h})((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}))|^{2+2\delta})^{\frac{1}{2+2\delta}} \\ &= \frac{1}{\sqrt{h}}(E|K(\frac{X_{t-1}-x}{h})(X_t - m(X_{t-1}))^2 \\ &\quad - E[K(\frac{X_{t-1}-x}{h})(X_t - m(X_{t-1}))^2|X_{t-1}]|^{2+2\delta})^{\frac{1}{2+2\delta}} \\ &\leq \frac{1}{\sqrt{h}}(E[K(\frac{X_{t-1}-x}{h})(X_t - m(X_{t-1}))^2]^{2+2\delta})^{\frac{1}{2+2\delta}} \\ &= \frac{1}{\sqrt{h}}(E[E[(K(\frac{X_{t-1}-x}{h})(X_t - m(X_{t-1}))^2]^{2+2\delta}|X_{t-1}]])^{\frac{1}{2+2\delta}} \\ &= \frac{1}{\sqrt{h}}(\int K^{2+2\delta}(v)\lambda(x+hv)p(x+hv)h dv)^{\frac{1}{2+2\delta}} \\ &= O(h^{\frac{1}{2+2\delta} - \frac{1+\delta}{2+2\delta}}) \leq O(h^{-\frac{\delta}{2+2\delta}}) = O(T^{\frac{1}{5} \frac{\delta}{2+\delta}}), \end{aligned}$$

wobei $\lambda(x) := E[(X_t - m(X_{t-1}))^{4+4\delta}|X_{t-1} = x] = \sigma^{4+4\delta}(x)E[\epsilon_t^{4+4\delta}|X_{t-1} = x]$.

Bei näherer Untersuchung des Beweises von Theorem 1 wird deutlich, dass sich die Aussage des Theorems für die uns vorliegende Situation auch unter der schwächeren Voraussetzung (i)' zeigen lässt: Die Beweisidee beruht auf einer Unterteilung der Summe der abhängigen Zufallsvariablen in alternierende lange und kurze Blöcke in der Art, dass die kurzen Blöcke einerseits klein genug sind, um nicht mehr zum Konvergenzverhalten der ursprünglichen Summe beizutragen, andererseits aber groß genug sind, um näherungsweise Unabhängigkeit der großen Blöcke untereinander zu garantieren. Dies ist eine klassische Vorgehensweise, um zentrale Grenzwertsätze für abhängige Zufallsvariablen zu beweisen, und wird gemeinhin Markov oder Bernstein (1927, Bernstein Summen) zugeschrieben.

Der derart modifizierte Satz lautet nun:

Theorem 2: Sei $\{X_{T,t} : 1 \leq t \leq d_T\}$ ein Dreiecksschema zentrierter Zufallsvariabler mit zugehörigen Mischkoeffizienten α_T . Gelten für ein $\delta > 0$:

(i) $\|X_{T,t}\|_{2+2\delta} = O(d_T^{\frac{1}{5}\frac{\delta}{2+\delta}})$ für alle T, t ,

(ii) $\sigma_{T,k_T}^2 := \text{Var} \frac{1}{\sqrt{k_T}} \sum_{t=1}^{k_T} X_{T,t} \rightarrow \tilde{\sigma}^2 > 0$ für alle Folgen $\{k_T\}$, die mit $T \rightarrow \infty$ gegen ∞ konvergieren,

(iii) $\alpha_T(k) \leq K\rho^k$, $0 \leq \rho < 1$

wobei K eine von T und k_T unabhängige Konstante ist,

so gilt:

$$\frac{1}{\sqrt{d_T}} \sum_{t=1}^{d_T} X_{T,t} \Longrightarrow \mathcal{N}(0, \tilde{\sigma}^2).$$

Anzumerken ist hier, dass statt Voraussetzung (iii) auch die schwächere Bedingung $\alpha_T(k) \leq K \cdot k^{-50}$ hinreichend ist.

Beweis: Definiere

$$U_{T,t} = X_{T,(t-1)(b_T+l_T)+1} + \dots + X_{T,(t-1)(b_T+l_T)+b_T}, \quad 1 \leq t \leq r_T,$$

wobei r_T die größte ganze Zahl i ist, für die $(t-1)(b_T+l_T)+b_T < d_T$. Definiere weiterhin

$$\begin{aligned} V_{T,t} &= X_{T,(t-1)(b_T+l_T)+b_T+1} + \dots + X_{T,i(b_T+l_T)+b_T}, \quad 1 \leq t < r_T, \\ V_{T,r_T} &= X_{T,(r_T-1)(b_T+l_T)+b_T+1} + \dots + X_{T,d_T}. \end{aligned}$$

Dann ist $S_{T,d_T,1} = \sum_{t=1}^{r_T} U_{T,t} + \sum_{t=1}^{r_T} V_{T,t}$. Nun ist l_T so zu wählen, dass l_T so klein ist, dass $\sum_{t=1}^{r_T} V_{T,t}$ klein im Vergleich zu $\sum_{t=1}^{r_T} U_{T,t}$ ist, und groß genug, um sicherzustellen, dass die $U_{T,t}$ fast unabhängig sind.

Wegen der Annahme der schwächeren Bedingung (i)' muss r_T hier größer und die Länge l_T der kurzen Blöcke kleiner gewählt werden als im ursprünglichen Beweis. Sei $b_T = \lfloor d_T^{\frac{3}{5}-\eta} \rfloor$ und $l_T = \lfloor d_T^{\frac{1}{6}-\mu} \rfloor$, $\eta > 0, \mu > 2\eta$, wobei $\lfloor \cdot \rfloor$ den ganzzahligen Anteil einer reellen Zahl bezeichnet.

Wie oben angemerkt, ist r_T die größte ganze Zahl i , für die $(t-1)(b_T+l_T)+b_T < d_T$ oder äquivalent $t-1 < \frac{d_T-b_T}{b_T+l_T}$ gilt, in unserem Fall gilt also

$$b_T \sim d_T^{\frac{3}{5}-\eta}, \quad l_T \sim d_T^{\frac{1}{6}-\mu}, \quad r_T \sim d_T^{\frac{2}{5}+\eta}. \quad (\text{iv})$$

Zunächst wird gezeigt, dass $d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} V_{T,t}$ für $d_T \rightarrow \infty$ nach Wahrscheinlichkeit gegen null konvergiert. Da sein Erwartungswert null ist, reicht es zu überprüfen, dass seine Varianz gegen null konvergiert. Nach Lemma A.0.3 aus Politis/Romano/Wolf (1999) und Voraussetzung (iii) gilt für alle T, t

$$E|V_{T,t}|^2 \leq 18K(\Delta d_T^{\frac{\delta}{10}})^2 l_T =: Bl_T d_T^{\frac{\delta}{5}} \quad (\text{v})$$

Also

$$\begin{aligned} \text{Var}^{\frac{1}{2}}(d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} V_{T,t}) &= E^{\frac{1}{2}}(d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} V_{T,t})^2 \\ &\leq \sum_{t=1}^{r_T} E^{\frac{1}{2}}(d_T^{-\frac{1}{2}} V_{T,t})^2 \\ &\leq \sum_{t=1}^{r_T-1} \left(\frac{Bl_T}{d_T^{1-\frac{\delta}{5}}} \right)^{\frac{1}{2}} + \left(\frac{B(l_T + b_T)}{d_T^{1-\frac{\delta}{5}}} \right)^{\frac{1}{2}} \quad (\text{wegen (v)}) \\ &\leq B^{\frac{1}{2}} \left(r_T \left(\frac{l_T}{d_T^{1-\frac{\delta}{5}}} \right)^{\frac{1}{2}} + \left(\frac{l_T + b_T}{d_T^{1-\frac{\delta}{5}}} \right)^{\frac{1}{2}} \right) \\ &= O(d_T^{\eta - \frac{1}{2}\mu + \frac{1}{10}\delta - \frac{1}{60}}) = o(1), \end{aligned}$$

falls $\delta < \frac{1}{6}$.

Nach dem Satz von Slutsky bleibt nun noch zu zeigen, dass $d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} U_{T,t} \Rightarrow \mathcal{N}(0, \tilde{\sigma}^2)$.

Seien $U'_{T,t}$, $1 \leq t \leq r_T$, unabhängige Zufallsvariablen mit denselben Verteilungen wie $U_{T,t}$, $1 \leq t \leq r_T$. Nach Lemma B.0.6 (Ibragimov) aus Politis/Romano/Wolf (1999) beträgt der Supremumsabstand zwischen den charakteristischen Funktionen von $d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} U_{T,t}$ und von $d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} U'_{T,t}$ höchstens $16r_T \alpha_T(l_T)$. Da die Mischkoeffizienten geometrisch fallend sind, r_T aber nur polynomial wachsend, ist dieser Ausdruck von der Ordnung $o(1)$. Daher ist der Beweis abgeschlossen, wenn gezeigt ist, dass

$$d_T^{-\frac{1}{2}} \sum_{t=1}^{r_T} U'_{T,t} \Rightarrow \mathcal{N}(0, \tilde{\sigma}^2).$$

Dieses wird in zwei Schritten nachgewiesen:

Schritt 1: $[Var(\sum_{t=1}^{r_T} U'_{T,t})]^{-\frac{2+\delta}{2}} \sum_{t=1}^{r_T} E|U'_{T,t}|^{2+\delta} \rightarrow 0$.

Schritt 2: $\frac{1}{d_T} Var(\sum_{t=1}^{r_T} U'_{T,t}) \rightarrow \tilde{\sigma}^2$.

Die Behauptung folgt dann durch Lyapunovs zentralen Grenzwertsatz und dem Satz von Slutsky.

Beweis von Schritt 1:

$$\begin{aligned}
\frac{1}{r_T b_T} Var(\sum_{t=1}^{r_T} U'_{T,t}) &= \frac{1}{r_T b_T} \sum_{t=1}^{r_T} E(U'_{T,t})^2 \\
&= \frac{1}{r_T b_T} \sum_{t=1}^{r_T} E(U_{T,t})^2 \\
&= \frac{1}{r_T} \sum_{t=1}^{r_T} E(b_T^{-\frac{1}{2}} U_{T,t})^2 \\
&= \frac{1}{r_T} \sum_{t=1}^{r_T} Var \frac{1}{\sqrt{b_T}} \sum_{t=1}^{b_T} X_{T,t} \\
&\rightarrow \tilde{\sigma}^2 \quad (\text{wegen(ii)}). \quad (\text{vi})
\end{aligned}$$

Nehme ohne Beschränkung der Allgemeinheit an, dass $1 \leq K$. Dann gilt nach Korollar A.0.1 aus Politis/Romano/Wolf (1999)

$$\begin{aligned}
E|b_T^{-\frac{1}{2}} U'_{T,t}|^{2+\delta} &\leq [3024 \max\{1, C_T^2(4)\}] 2^{4(4(2-\delta)/\delta+1)} (\Delta \cdot d_T^{\frac{1}{5} \frac{\delta}{2+\delta}})^{(2+\delta)(1+\frac{\delta}{2})} \\
&\leq [3024 K^2] 2^{4(4(2-\delta)/\delta+1)} (\Delta' \cdot d_T^{\frac{1}{5} \delta})^{(1+\frac{\delta}{2})} \quad (\text{wegen(iii)}) \\
&=: D_T \quad (\text{vii})
\end{aligned}$$

Schließlich gilt

$$\begin{aligned}
& [Var(\sum_{t=1}^{r_T} U'_{T,t})]^{-\frac{2+\delta}{2}} \sum_{t=1}^{r_T} E|U'_{T,t}|^{2+\delta} \\
&= \left(\frac{1}{r_T}\right)^{\frac{2+\delta}{2}} \left[\frac{1}{r_T b_T} Var(\sum_{t=1}^{r_T} U'_{T,t})\right]^{-\frac{2+\delta}{2}} \sum_{t=1}^{r_T} E|b_T^{-\frac{1}{2}} U_{T,t}|^{2+\delta} \\
&\leq \left(\frac{1}{r_T}\right)^{\frac{2+\delta}{2}} \left[\frac{1}{r_T b_T} Var(\sum_{t=1}^{r_T} U'_{T,t})\right]^{\frac{2}{2+\delta}} r_T D_T \quad \text{wegen (vii)} \\
&\leq \left(\frac{1}{r_T}\right)^{\frac{2+\delta}{2}} O(1) O(r_T) O(d_T^{\frac{1}{5} \frac{\delta}{2+\delta} (2+\delta)(1+\frac{\delta}{2})}) \quad \text{wegen (vi)} \\
&\leq O(r_T^{-\frac{\delta}{2}}) O(d_T^{\frac{1}{5} \delta (1+\frac{\delta}{2})}) \\
&\leq O(d_T^{-\frac{\delta}{2} (\frac{2}{5} + \eta) + \frac{1}{5} \delta (1+\frac{\delta}{2})}) \\
&\leq O(d_T^{-\frac{\delta}{2} (\eta - \frac{\delta}{5})}) = o(1)
\end{aligned}$$

für $\eta > \frac{\delta}{5}$.

Beweis von Schritt 2: Wir haben gerade gezeigt, dass $\frac{1}{r_T b_T} Var(\sum_{t=1}^{r_T} U'_{T,t}) \rightarrow \tilde{\sigma}^2$. Nach (iv) gilt außerdem $\frac{r_T b_T}{d_T} \rightarrow 1$. \square

Damit bleibt nur noch Bedingung (ii) zu überprüfen

$$\begin{aligned}
& \sigma_{T,k_T}^2 \\
&= Var\left(\frac{1}{\sqrt{k_T}} \sum_{t=1}^{k_T} X_{T,t}\right) \\
&= \frac{1}{k_T} \sum_{t=1}^{k_T} Var\left(\frac{1}{\sqrt{h}} K\left(\frac{X_{t-1}-x}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}))\right) \\
&\quad + \sum_{s < t} Cov\left(\frac{1}{\sqrt{h}} K\left(\frac{X_{s-1}-x}{h}\right) ((X_s - m(X_{s-1}))^2 - \sigma^2(X_{s-1})), \right. \\
&\quad \left. \frac{1}{\sqrt{h}} K\left(\frac{X_{t-1}-x}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}))\right) \\
&= \frac{1}{k_T} \sum_{t=1}^{k_T} E \frac{1}{h} K^2\left(\frac{X_{t-1}-x}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}))^2
\end{aligned}$$

$$\begin{aligned}
& + \frac{2}{k_t} \sum_{t=1}^{k_T} \sum_{s=1}^{t-1} Cov\left(\frac{1}{\sqrt{h}} K\left(\frac{X_{s-1}-x}{h}\right) ((X_s - m(X_{s-1}))^2 - \sigma^2(X_{s-1})) , \right. \\
& \left. \frac{1}{\sqrt{h}} K\left(\frac{X_{t-1}-x}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1})) \right) \\
\leq & E \frac{1}{h} K^2\left(\frac{X_0-x}{h}\right) ((X_1 - m(X_0))^2 - \sigma^2(X_0))^2 \\
& + \frac{2}{h} \sum_{k=1}^{\infty} Cov\left(K\left(\frac{X_0-x}{h}\right) ((X_1 - m(X_0))^2 - \sigma^2(X_0)), \right. \\
& \left. K\left(\frac{X_k-x}{h}\right) ((X_{k+1} - m(X_k))^2 - \sigma^2(X_k))\right)
\end{aligned}$$

Betrachte nun den ersten Summanden

$$\begin{aligned}
& E \frac{1}{h} K^2\left(\frac{X_0-x}{h}\right) ((X_1 - m(X_0))^2 - \sigma^2(X_0))^2 \\
& = E[E[\frac{1}{h} K^2\left(\frac{X_0-x}{h}\right) (\sigma^2(X_0) (\frac{X_1 - m(X_0)}{\sigma(X_0)})^2 - 1)^2 | X_0]] \\
& = E[\frac{1}{h} K^2\left(\frac{X_0-x}{h}\right) \sigma^4(X_0) E[(\epsilon_1^2 - 1)^2 | X_0]] \\
& = E[\frac{1}{h} K^2\left(\frac{X_0-x}{h}\right) \sigma^4(X_0) (E[\epsilon_1^4 | X_0] - 1)] \\
& = \frac{1}{h} \int K^2\left(\frac{u-x}{h}\right) \sigma^4(u) (\kappa(u) - 1) p(u) du \\
& = \frac{1}{h} \int K^2(v) \sigma^4(x - hv) [\kappa(x - hv) - 1] p(x - hv) h dv \\
& \rightarrow \int K^2(v) dv \sigma^4(x) p(x) (\kappa(x) - 1)
\end{aligned}$$

Der zweite Summand verhält sich wie folgt

$$\begin{aligned}
& \frac{2}{h} \sum_{k=1}^{\infty} Cov(K(\frac{X_0 - x}{h})((X_1 - m(X_0))^2 - \sigma^2(X_0))) \\
& \quad K(\frac{X_k - x}{h})((X_{k+1} - m(X_k))^2 - \sigma^2(X_k))) \\
&= \frac{2}{h} \sum_{k=1}^M EK(\frac{X_0 - x}{h})K(\frac{X_k - x}{h})((X_1 - m(X_0))^2 - \sigma^2(X_0)) \\
& \quad ((X_{k+1} - m(X_k))^2 - \sigma^2(X_k)) \\
& \quad + \frac{2}{h} \sum_{k=M+1}^{\infty} Cov(K(\frac{X_0 - x}{h})((X_1 - m(X_0))^2 - \sigma^2(X_0))) \\
& \quad K(\frac{X_k - x}{h})((X_{k+1} - m(X_k))^2 - \sigma^2(X_k))) \\
&\leq \frac{2}{h} c_1 c_2 \sum_{k=1}^M EK(\frac{X_0 - x}{h})K(\frac{X_k - x}{h}) \\
& \quad + \frac{2}{h} \sum_{k=M+1}^{\infty} \alpha_k^{\frac{\delta}{2+\delta}} \cdot c_3 \cdot (E|K(\frac{X_0 - x}{h})((X_1 - m(X_0))^2 - \sigma^2(X_0))|^{2+\delta})^{\frac{2}{2+\delta}} \\
&= \frac{2}{h} c_1 c_2 \sum_{k=1}^M h^2 \int \int K(v_1)K(v_2)p(x + hv_1)p(x + hv_2)dv_1 dv_2 \\
& \quad + \frac{2}{h} c_3 \sum_{k=M+1}^{\infty} \alpha_k^{\frac{\delta}{2+\delta}} \underbrace{(EK^{2+\delta}(\frac{X_0 - x}{h})|(X_1 - m(X_0))^2 - \sigma^2(X_0)|^{2+\delta})^{\frac{2}{2+\delta}}}_{\leq h^{\frac{2}{2+\delta}} \cdot c_4 = h^{1 - \frac{\delta}{2+\delta}} \cdot c_4} \\
&\leq c \cdot M \cdot h + c' \cdot \frac{1}{h} h^{1-\epsilon} \sum_{k=M+1}^{\infty} (\rho^{\frac{\delta}{2+\delta}})^k, \quad 0 < \epsilon := \frac{\delta}{2+\delta} < 1 \\
&= c \cdot M \cdot h + c' \cdot h^{-\epsilon} \sum_{k=M+1}^{\infty} \tilde{\rho}^k \quad \text{Wähle } M = \ln T^\gamma. \\
&\leq o(1) + h^{-\epsilon} \tilde{\rho}^{\ln T^\gamma} \cdot c'' \quad \text{Setze } \tilde{\rho} = e^{-r} \text{ für ein } r > 0. \\
&= o(1) + T^{\frac{\epsilon}{5}} \cdot e^{-\ln T^{r\gamma}} \cdot c'' = o(1) + T^{\frac{\epsilon}{5} - r\gamma} = o(1)
\end{aligned}$$

für $\gamma > \frac{\epsilon}{5r}$.

Die Asymptotik des Biasanteils lässt sich genauso wie im modellabhängigen Fall nachrechnen.

Der Beweis des asymptotischen Verschwindens der übrigen Terme verläuft analog zum modellabhängigen Fall, wobei jedoch zu beachten ist, dass die bedingten Erwartungswerte nun nur unter dem direkten Vorgänger X_{t-1} und nicht unter der gesamten Vergangenheit \mathcal{F}_{t-1} bedingt sind. Dabei lässt sich die Burkholder-Ungleichung nicht anwenden, da kein Martingal mehr vorliegt. Hier liefert aber mit entsprechenden Umformungen wie im modellabhängigen Fall die Ungleichung aus Doukhan die geforderte Konvergenz gegen null.

Darüber hinaus ist es nun nicht mehr möglich die Modellgleichung (1.1) zu verwenden. Allerdings lassen sich die betroffenen Terme durch andere Umformungen, nämlich das Einschieben eines unter dem unmittelbaren Vorgänger bedingten Erwartungswertes in die Erwartung (siehe Beginn dieses Kapitels) auf die entsprechende Form im modellabhängigen Fall bringen.

Die asymptotische Normalität des Nadaraya-Watson-Schätzers der bedingten Varianz ist also auch ohne einen zugrundeliegenden NARCH-Prozess gezeigt worden. Im folgenden geben wir noch einige Erklärungen zum Vorgehen beim Wild Bootstrap ohne zugrundeliegenden NARCH-Prozess. Im wesentlichen gehen wir genauso vor wie in Kapitel 3.1.

Definiere

$$\begin{aligned}\eta_t &:= \sigma^2(X_{t-1})(\epsilon_t^2 - 1) \\ &= (X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}),\end{aligned}$$

wobei $\epsilon_t := \frac{X_t - m(X_{t-1})}{\sigma(X_{t-1})}$. Es gilt offenbar $E\eta_t = 0$, $Var\eta_t = E[\sigma^2(X_{t-1})(\epsilon_t^2 - 1)]^2 = E(\sigma^4(X_{t-1})(E[\epsilon_t^4|X_{t-1}] - 1))$. Berechne die empirischen Residuen.

$$\hat{\eta}_t := \hat{\sigma}_h^2(X_{t-1})(\hat{\epsilon}_t^2 - 1) = (X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1}),$$

wobei $\hat{\epsilon}_t := \frac{X_t - \hat{m}_h(X_{t-1})}{\hat{\sigma}_h(X_{t-1})}$. Erzeuge ein unabhängiges Resample $\{\eta_t^*\}$, wobei für den (bedingten) Erwartungswert $E^*\eta_t^* = 0$ und für die Varianz $E^*\eta_t^{*2} = \hat{\eta}_t^2$ gelte. Dabei bezeichne E^* den bedingten Erwartungswert, gegeben X_0, \dots, X_T . Dieses ist beispielsweise zu erreichen durch Erzeugung einer unabhängig identisch verteilten Stichprobe (Sample) $\bar{\eta}_1, \dots, \bar{\eta}_T$ mit Erwartungswert null und Varianz eins und Definition von $\eta_t^* = \hat{\eta}_t \bar{\eta}_t$. Zusätzlich werde ein endliches viertes Moment von $\bar{\eta}_t$ vorausgesetzt.

Die Wild Bootstrap-Stichprobe (Resample) wird wie folgt definiert

$$X_t^{2*} = \hat{\sigma}_g^2(X_{t-1}) + \eta_t^*. \quad (3.3)$$

Das lässt sich in Analogie zur vorhergehenden Notation auch schreiben als:

$$\eta_t^* = \hat{\sigma}_g^2(X_{t-1}) \left(\frac{X_t^{2*}}{\hat{\sigma}_g^2(X_{t-1})} - 1 \right).$$

Den Bootstrap-Schätzer für die Varianz σ^2 definieren wir wie in Kapitel 3.1:

$$\hat{\sigma}_h^{*2}(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right) X_t^{2*}}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right)}. \quad (3.4)$$

Lemma 3.6: Gelten (W1)-(W7). Dann gilt für ein $x \in \mathbb{R}$

$$\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x-X_{t-1}}{h}\right) \eta_t^* \implies \mathcal{N}(0, \tau^2(x)) \quad \text{nach Wahrscheinlichkeit,}$$

wobei $\tau^2(x) = \sigma^4(x)p(x) \int K^2(v)dv (E[\epsilon_1^4|X_0 = x] - 1)$.

Der Beweis läuft analog zum modellabhängigen Fall, nur dass hier nicht mehr die Modellvorstellung verwendet werden kann.

Für die Varianz z.B. erhalten wir aber

$$\begin{aligned} & E^* \left(\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x-X_{t-1}}{h}\right) \eta_t^* \right)^2 \\ &= \frac{1}{Th} \sum_t K^2\left(\frac{x-X_{t-1}}{h}\right) (\sigma^2(X_{t-1})(\epsilon_t^2 - 1))^2 + o_P(1) \\ &\rightarrow \sigma^4(x)p(x) \int K^2(v)dv (E[\epsilon_1^4|X_0 = x] - 1) \quad \text{nach Wahrscheinlichkeit} \end{aligned}$$

denn:

$$\begin{aligned} & E\left[\frac{1}{Th} \sum_t K^2\left(\frac{x-X_{t-1}}{h}\right) (\sigma^2(X_{t-1})(\epsilon_t^2 - 1))^2\right] \\ &= \frac{1}{h} E\left[E\left[K^2\left(\frac{x-X_0}{h}\right) (\sigma^2(X_0)(\epsilon_1^2 - 1))^2 \middle| X_0\right]\right] \\ &= \frac{1}{h} E\left[K^2\left(\frac{x-X_0}{h}\right) \sigma^4(X_0) E[(\epsilon_1^2 - 1)^2 | X_0]\right] \\ &= \sigma^4(x)p(x) \int K^2(v)dv (E[\epsilon_1^4|X_0 = x] - 1) \end{aligned}$$

und

$$Var[\frac{1}{Th} \sum_t K^2(\frac{x - X_{t-1}}{h}) (\sigma^2(X_{t-1})(\epsilon_t^2 - 1))^2] = o(1)$$

genau wie im modellabhängigen Fall.

Kapitel 4

Simulationen

In diesem Abschnitt untersuchen wir das Verhalten des Wild Bootstrap bei endlichen Stichprobengrößen. Um die Güte der Approximation durch Wild Bootstrap zu bestimmen, betrachten wir simulierte Realisierungen des folgenden Prozesses

$$X_t = 0.4 \cdot X_{t-1} + \sqrt{0.5 + 0.25 \cdot X_{t-1}^2} \cdot \epsilon_t, \quad t = 1, \dots, T,$$

wobei die Zufallsvariablen ϵ_t , $t = 1, \dots, T$ unabhängig identisch standard-normalverteilt angenommen werden und die Stichprobengröße auf $T = 100$, 200 bzw. 500 gesetzt wird. Die Dichte der Verteilung von

$$\sqrt{Th}(\hat{\sigma}_h^2(0) - \sigma^2(0))/\hat{V}_h(0),$$

wobei

$$\hat{V}_h^2(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x-X_{t-1}}{h}\right) ((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(X_{t-1}))^2}{\left(\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x-X_{t-1}}{h}\right)\right)^2},$$

wird aus 1000 Wiederholungen einer Monte-Carlo-Simulation berechnet. Die Standardisierung der Verteilung stellt sicher, dass die Verteilung asymptotisch normalverteilt ist mit Varianz 1 und i.a. nicht verschwindendem Erwartungswert.

Die Wild-Bootstrap-Approximation ist gegeben durch die Verteilung von

$$\sqrt{Th}(\hat{\sigma}_h^{*2}(0) - \hat{\sigma}_g^2(0))/\hat{V}_h(0),$$

wobei wir die gleiche Standardisierung wie oben benutzen. Wir verwenden einen so genannten überglätteten Kernschätzer $\hat{\sigma}_g^2$ mit einer Bandweite $g > h$. Als Faustregel wird häufig $g \approx 1.5h$ oder $g \approx 2h$ vorgeschlagen. Jede Bootstrap-Verteilung, die natürlich eine bedingte Verteilung, gegeben X_0, \dots, X_T ist, wurde aus einer Simulation mit 1000 Wild-Bootstrap-Stichprobenwiederholungen gewonnen. Die Qualität der Wild-Bootstrap-Approximation wird nicht nur anhand einer Realisation einer Zeitreihe X_1, \dots, X_T beurteilt. Es wurden 40 Zeitreihen der Länge T erzeugt und für jede wurde die Wild-Bootstrap-Verteilung berechnet. Der Übersichtlichkeit halber werden aber nicht alle 40, sondern nur drei repräsentative Fälle gezeigt. Für jede der 40 Bootstrap-Dichten $f_i^*, i = 1, \dots, 40$ haben wir die mittlere quadratische Abweichung, den MSE (Mean Square Error), zur wahren Dichte f berechnet, d.i. $MSE_i = \int (f_i^*(x) - f(x))^2 dx$. Wir zeigen die Wild-Bootstrap-Dichte, die dem Median der MSE-Werte entspricht, als durchschnittlichen (gelbe Kurve), jene, welche dem oberen Quartil der MSE-Werte entspricht, als unterdurchschnittlichen (blaue Kurve), und schließlich jene, welche dem unteren Quartil der MSE-Werte entspricht, als überdurchschnittlichen Fall (rote Kurve).

Abbildung 4.1 enthält diese drei aus 100 Wild-Bootstrap-Dichten zusammen mit der zugrunde liegenden wahren Dichte (grüne Kurve), die approximiert werden soll, für die Stichprobengrößen T und Bandweiten g und h mit $T = 100$, $h = 0.8$ und $g = 1.1$, während die Abbildungen 4.2 und 4.3 die entsprechenden Ergebnisse für $T = 200$, $h = 0.65$ und $g = 0.9$ bzw. $T = 500$, $h = 0.5$ und $g = 0.7$ zeigen.

Abbildung 4.4 zeigt die täglichen Log>Returns des DAX vom 1. Januar 1990 bis zum 31. Dezember 1992, Abbildung 4.5 den lokal-konstanten Nadaraya-Watson-Schätzer der Volatilität des DAX in diesem Zeitraum. Die täglichen Log-returns des Wechselkurses US-Dollar - D-Mark vom 1. Januar 1989 bis zum 31. Dezember 1998 sind in Abbildung 4.6 dargestellt. Abbildung 4.7 stellt den lokal-konstanten Nadaraya-Watson-Schätzer der Volatilität des Wechselkurses US-Dollar - D-Mark dar.

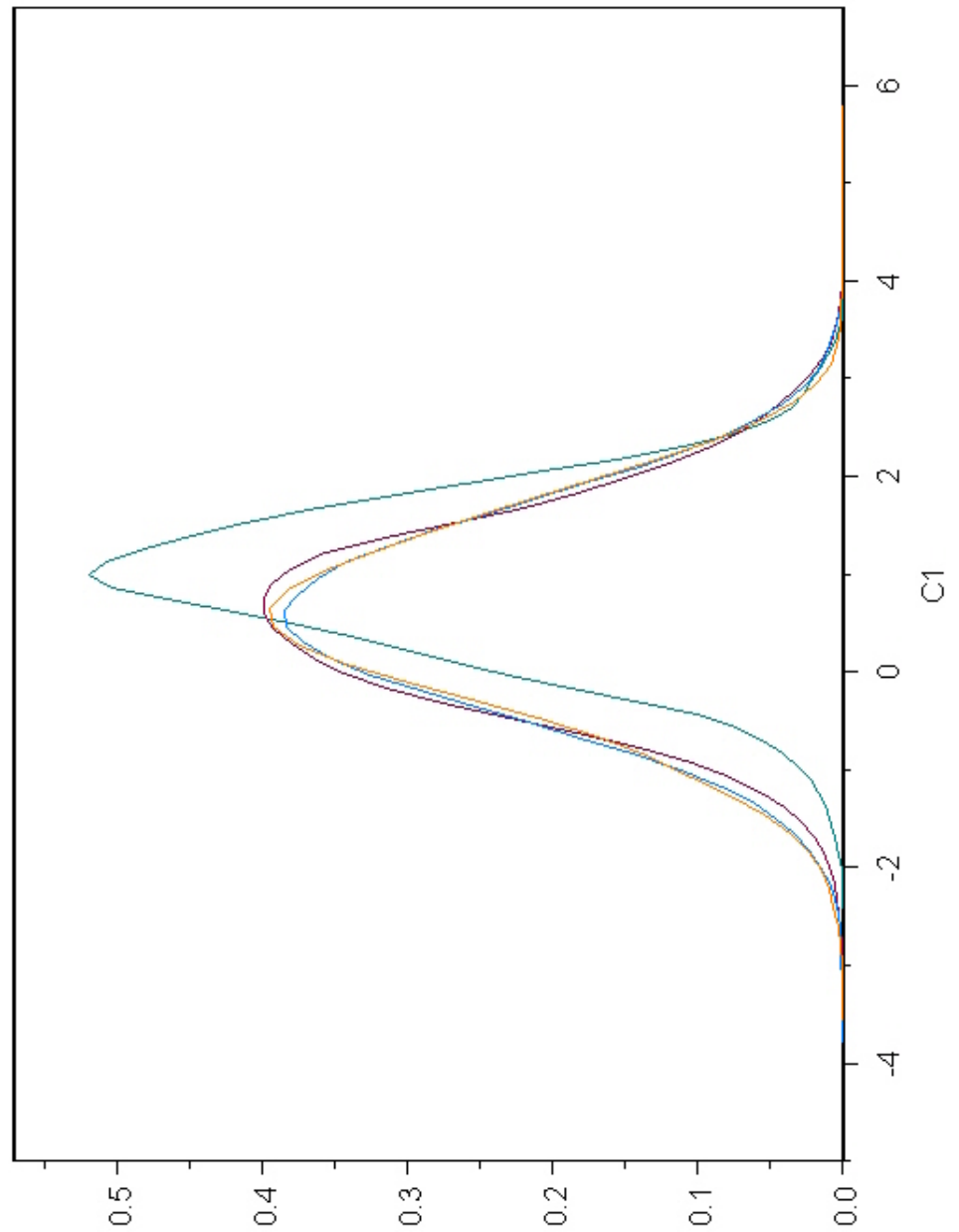


Abbildung 4.1: Nadaraya-Watson - Wild Bootstrap, $T=100$

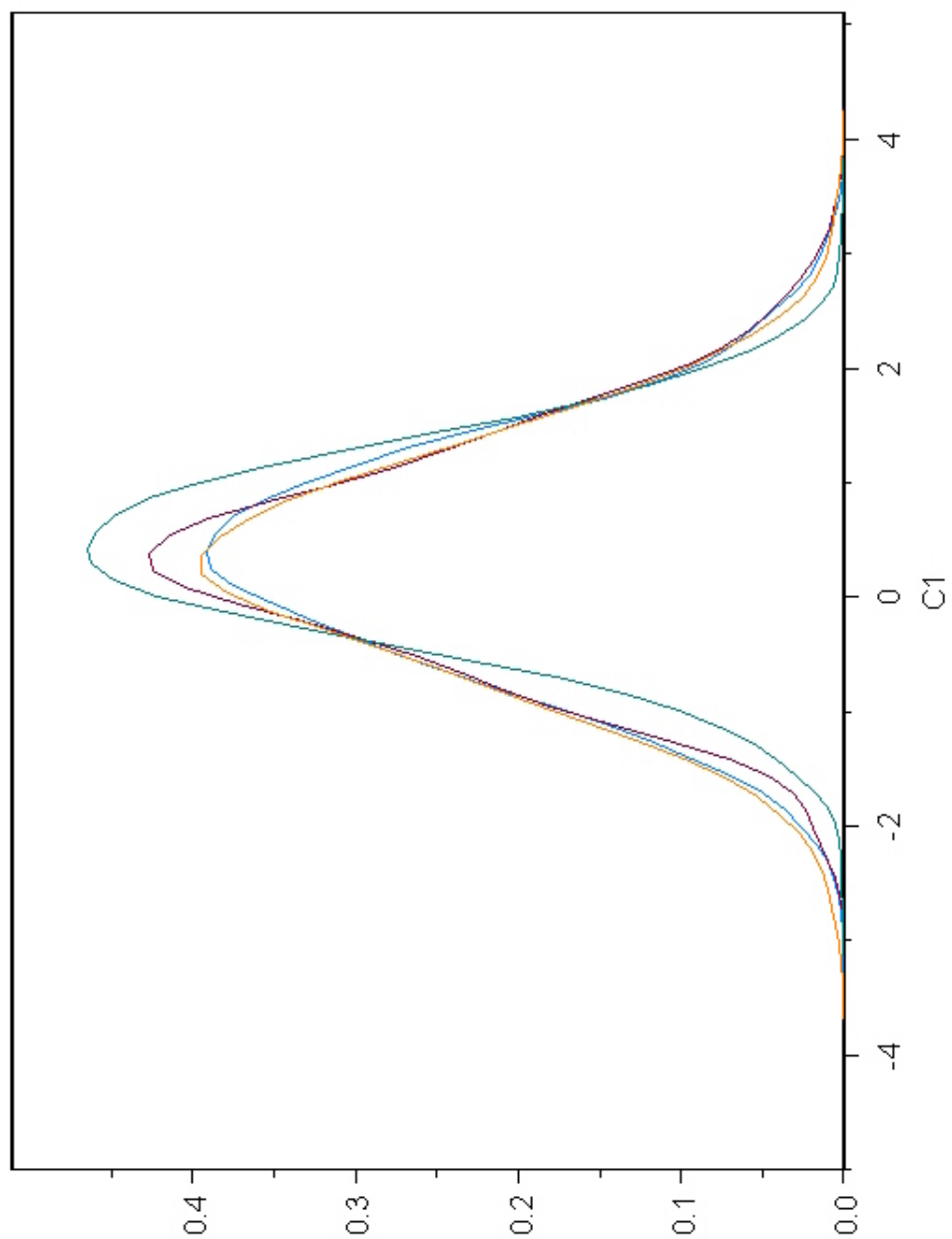


Abbildung 4.2: Nadaraya-Watson - Wild Bootstrap, $T=200$

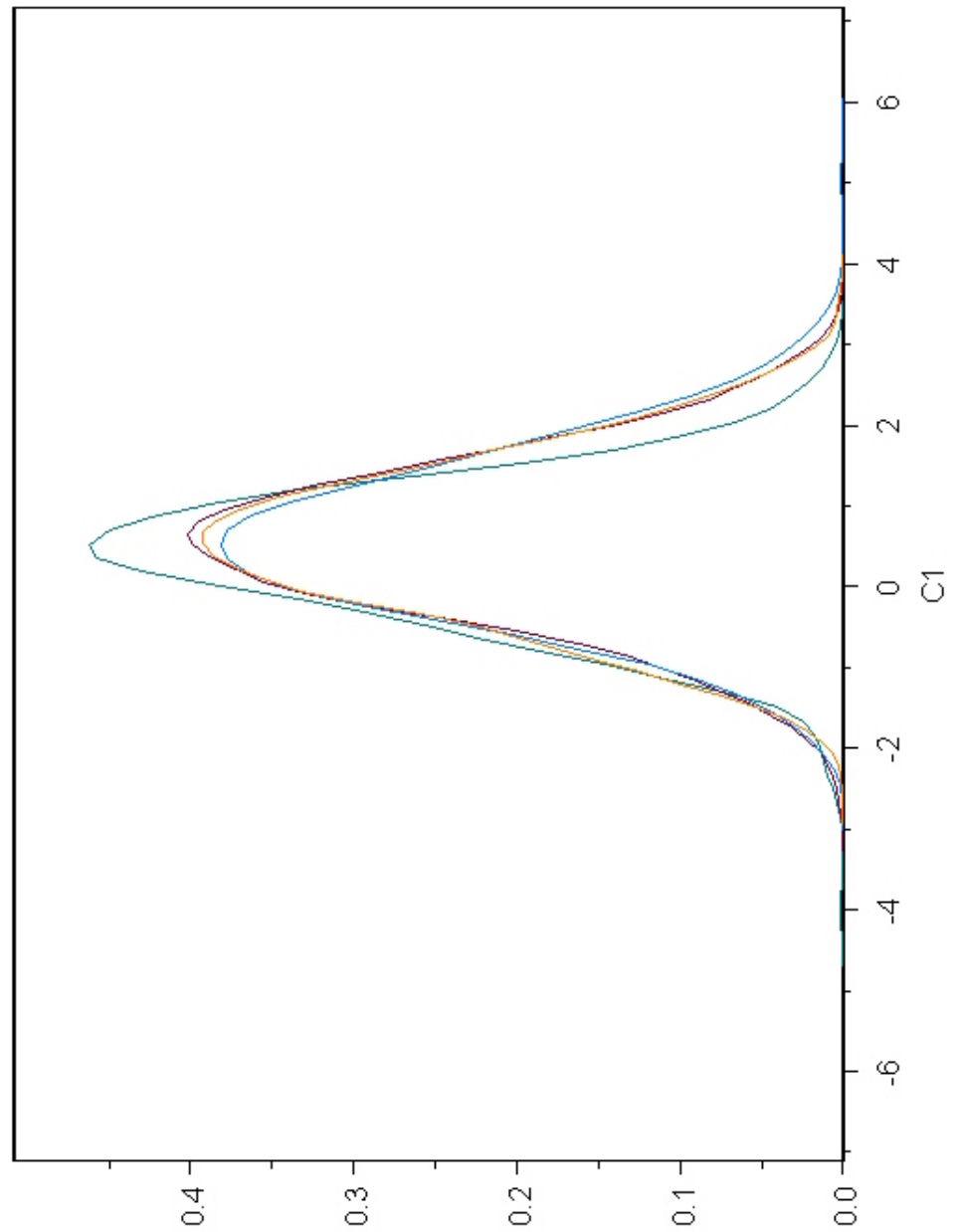


Abbildung 4.3: Nadaraya-Watson - Wild Bootstrap, $T=500$

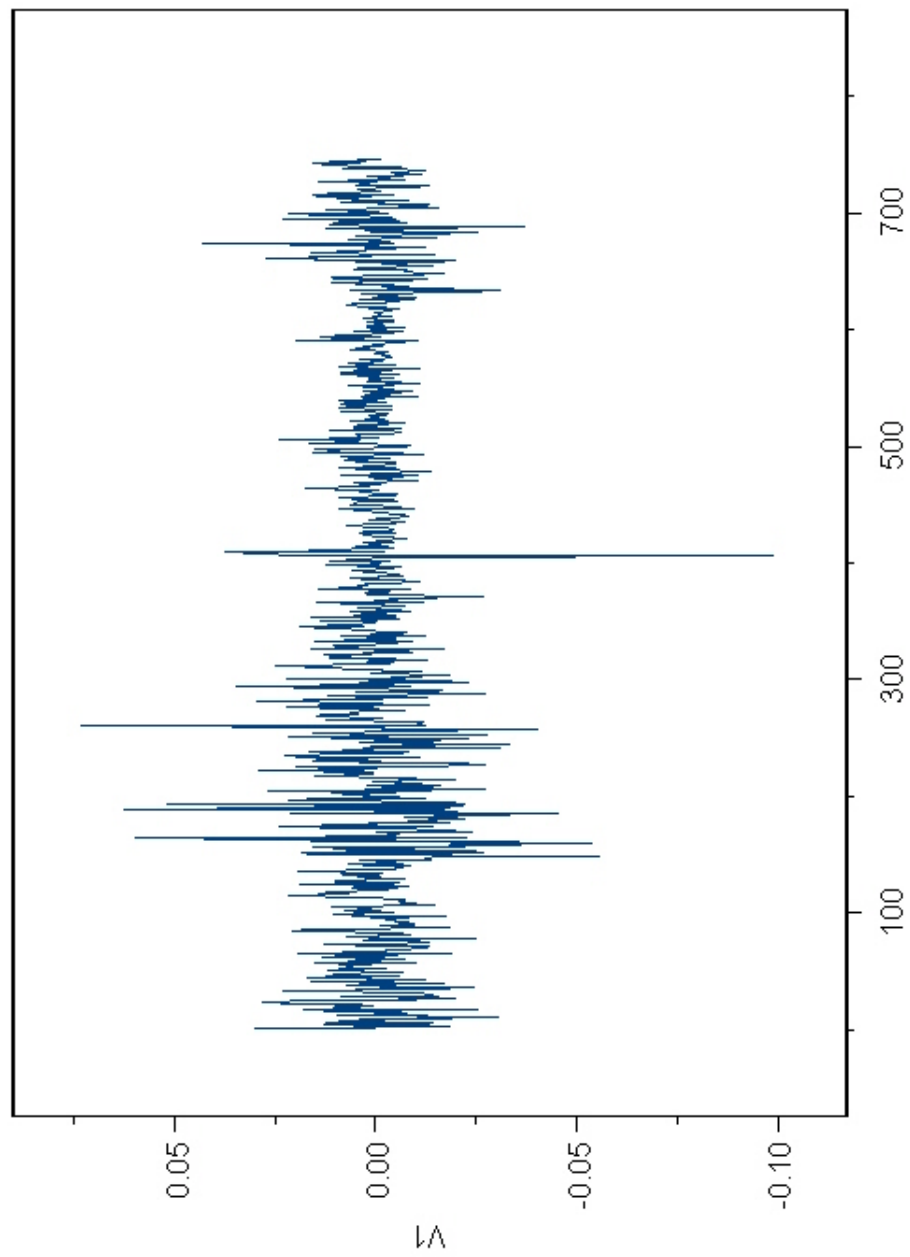


Abbildung 4.4: Tägliche Log-returns des DAX (1990-92)

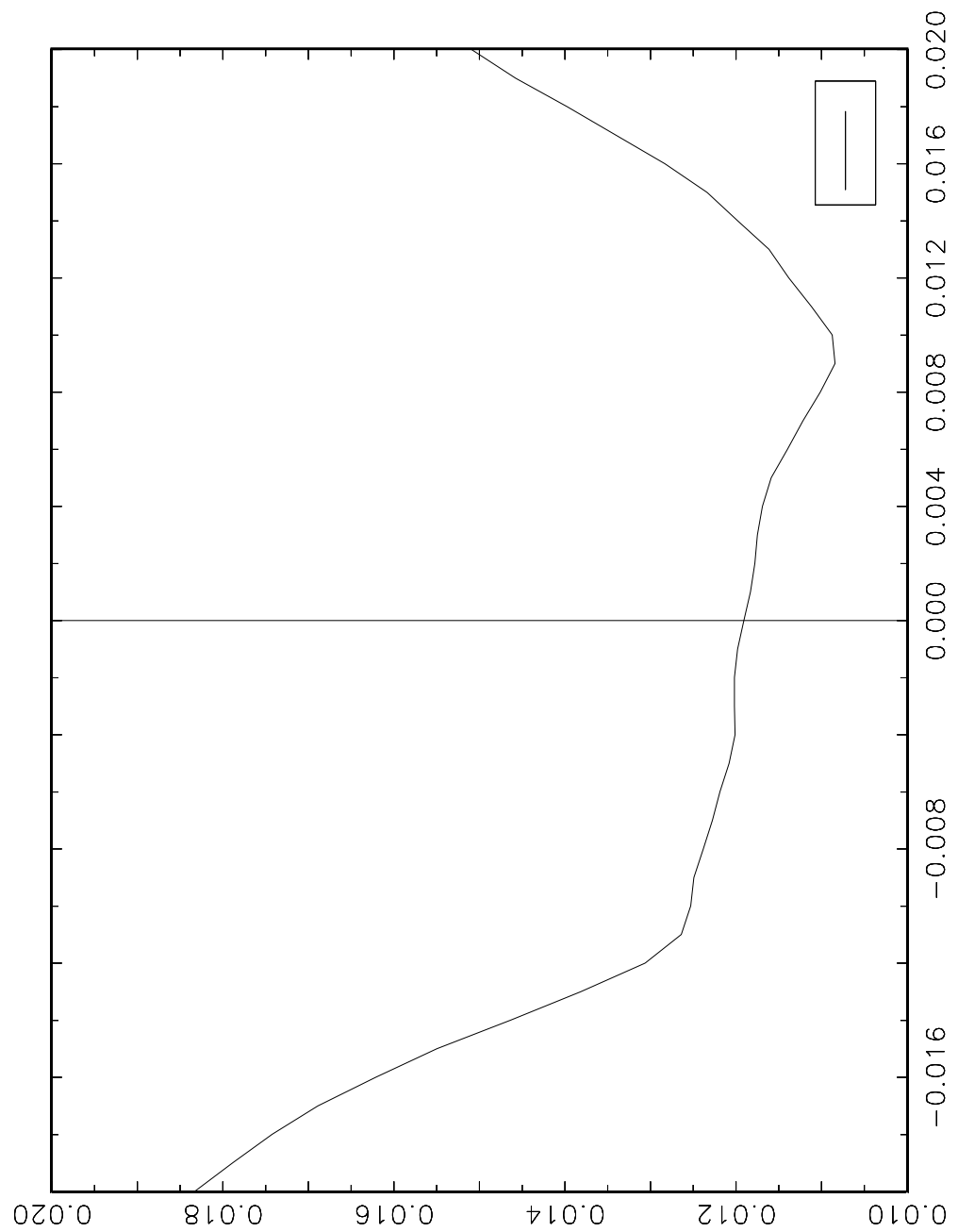


Abbildung 4.5: Lokal-konstanter Schätzer der Volatilität des DAX (1990-92)

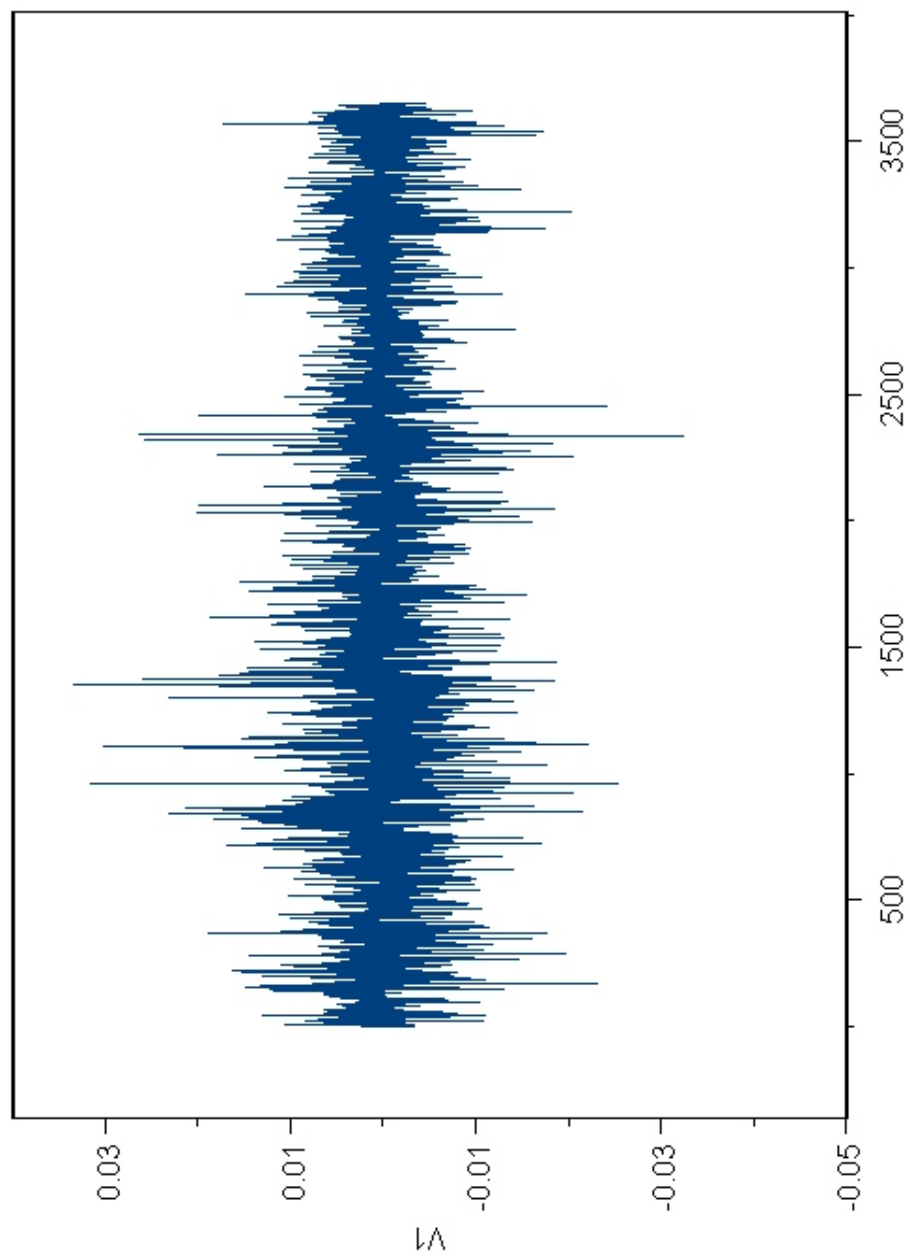


Abbildung 4.6: Tägliche Log-returns des Wechselkurses US \$ - DM (1989-98)

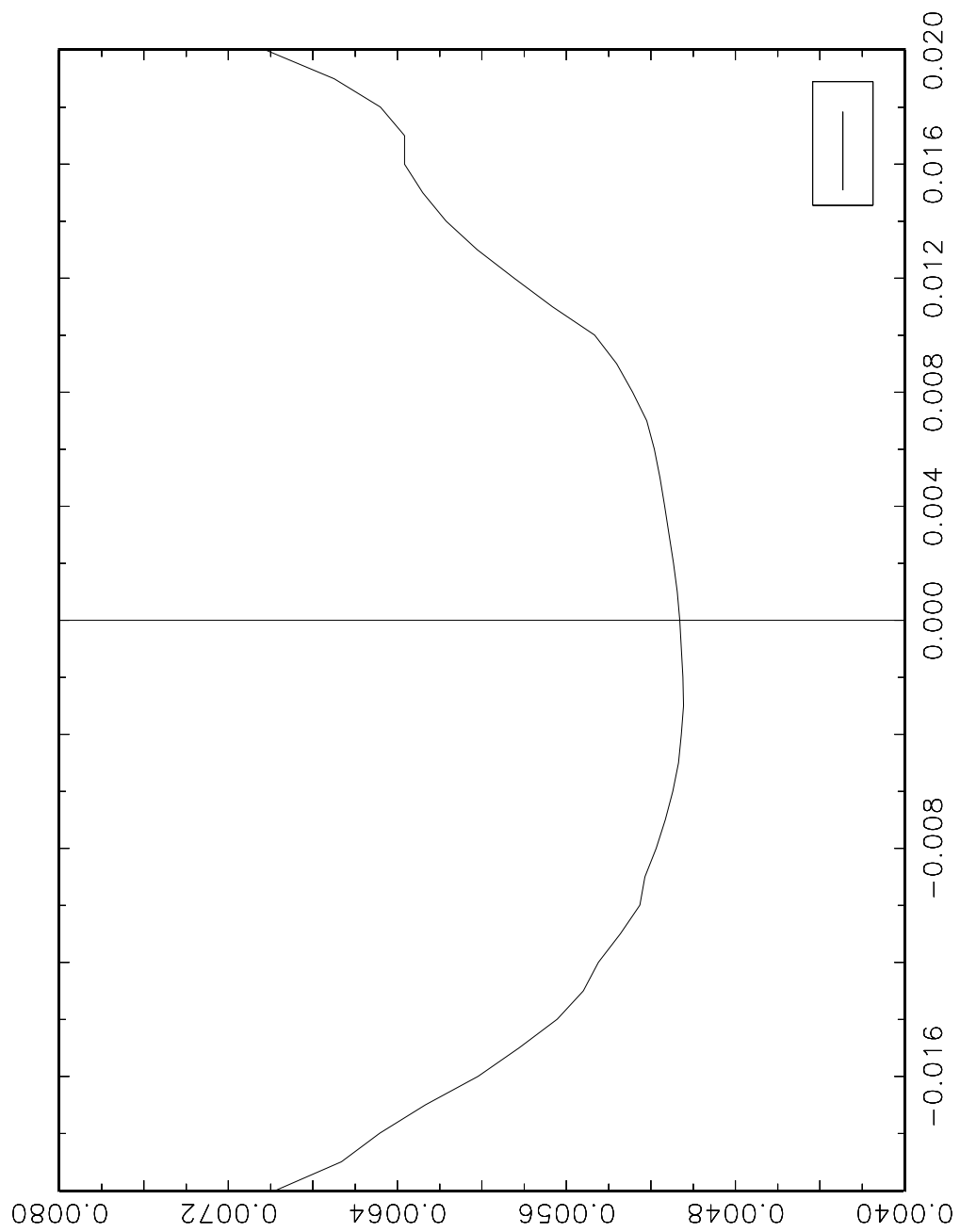


Abbildung 4.7: Lokal-konstanter Schätzer der Volatilität des Wechselkurses US \$ - DM (1989-98)

Kapitel 5

Paarweises Bootstrap

Es liegen die Beobachtungspaare $(X_0, X_1), (X_1, X_2), \dots, (X_{T-1}, X_T)$ des CHARN-Modells (2.1) bzw. allgemeiner eines beliebigen stark mischenden und stationären Prozesses vor. Aus ihnen wird nun T -mal unabhängig mit Zurücklegen gezogen. So erhält man die Bootstrap-Stichprobe $(X_0^*, X_1^*), (X_1^*, X_2^*), \dots, (X_{T-1}^*, X_T^*)$. Da die Bootstrap-Stichproben gemäß einer Laplace-Verteilung aus den ursprünglichen Beobachtungspaaren gezogen werden, gilt $P^*\{(X_{t-1}^*, X_t^*) = (X_{s-1}, X_s)\} = \frac{1}{T}$ für alle $s, t = 1, \dots, T$, wobei P^* die bedingte Verteilung von (X_{t-1}^*, X_t^*) , gegeben $(X_0, X_1), \dots, (X_{T-1}, X_T)$ bezeichne.

Durch dieses Vorgehen wird erreicht, dass im Gegensatz zum Wild Bootstrap die Abhängigkeit zwischen zwei (zeitlich) aufeinanderfolgenden Zufallsvariablen des Prozesses erhalten bleibt.

Auf der Grundlage der Bootstrap-Stichprobenpaare $(X_0^*, X_1^*), \dots, (X_{T-1}^*, X_T^*)$ kann man analog zu Kapitel 1.1 einen lokal-konstanten Kernschätzer definieren

$$\hat{\sigma}_h^{*2}(x) = \frac{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right) (X_t^* - \hat{m}_h(X_{t-1}^*))^2}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right)}. \quad (5.1)$$

Im folgenden soll gezeigt werden, dass dieser Schätzer dieselbe asymptotische Verteilung wie der Wild Bootstrap-Schätzer von σ besitzt, d.h. dass also auch das paarweise Bootstrap eine konsistente Schätzung von $\mathcal{L}^\sigma(x)$ erlaubt.

Satz 5.1: Es gelten die Voraussetzungen (W1)-(W7) aus dem vorigen Abschnitt und für die Bootstrap-Stichprobenpaare: $P^*\{(X_{t-1}^*, X_t^*) = (X_{s-1}, X_s)\} = \frac{1}{T}$ für alle $s, t = 1, \dots, T$.

Dann gilt für ein $x \in \mathbb{R}$:

$$\sqrt{Th}(\hat{\sigma}_h^{*2}(x) - \hat{\sigma}_h^2(x)) \implies \mathcal{N}(0, \tau^2(x)) \quad \text{nach Wahrscheinlichkeit,}$$

wobei

$$\tau^2(x) = \frac{\sigma^4(x)}{p(x)} \int K^2(v) dv (\kappa(x) - 1)$$

Beweis: Einsetzen von $\hat{\sigma}_h^{*2}(x)$ ergibt

$$\sqrt{Th}(\hat{\sigma}_h^{*2}(x) - \hat{\sigma}_h^2(x)) = \frac{\frac{1}{\sqrt{Th}} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))}{\frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right)}.$$

Berechne Erwartungswert des Nenners:

$$\begin{aligned} E^* \frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right) &= \frac{1}{h} E^* K\left(\frac{x - X_1^*}{h}\right) \\ &= \frac{1}{Th} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h}\right) \\ &\rightarrow p(x) \end{aligned}$$

nach Wahrscheinlichkeit. Zeige Konvergenz gegen den zugehörigen Erwartungswert:

$$\begin{aligned} &E^* \left(\frac{1}{Th} \sum_{t=1}^T \left(K\left(\frac{x - X_{t-1}^*}{h}\right) - E^* K\left(\frac{x - X_{t-1}^*}{h}\right) \right)^2 \right) \\ &= \frac{1}{Th^2} E^* \left(K\left(\frac{x - X_{t-1}^*}{h}\right) - E^* K\left(\frac{x - X_{t-1}^*}{h}\right) \right)^2 \\ &\leq \frac{1}{Th^2} E^* K^2\left(\frac{x - X_{t-1}^*}{h}\right) \\ &= \frac{1}{T^2 h^2} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}^*}{h}\right) \\ &= O\left(\frac{1}{Th}\right) = o(1) \end{aligned}$$

Das Konvergenzverhalten des Zählers lässt sich mit Hilfe des Grenzwertsatzes von Lindeberg für Dreiecksschemata untersuchen, da die einzelnen Summanden wegen des unabhängigen Ziehens mit Zurücklegen voneinander unabhängig sind.

Das Dreiecksschema lautet in diesem Fall

$$X_{T,t} := \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))$$

Folgende Bedingungen müssen erfüllt sein:

$$(i) \ E^* X_{T,t} = 0$$

$$(ii) \ Var^* X_{T,t} < \infty$$

$$(iii) \ \frac{1}{s_n^2} \sum_{t=1}^T \int 1_{\{|X_{T,t}| \geq \epsilon s_n\}} X_{T,t}^2 dP^* \rightarrow 0 \text{ für alle } \epsilon > 0$$

$$\text{wobei } s_n^2 = \sum_{t=1}^T E^* X_{T,t}^2 \text{ (Lindebergbedingung)}$$

zu (i):

$$\begin{aligned} E^* \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x)) \\ = \frac{1}{\sqrt{Th}} \frac{1}{T} \sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right) [(X_t - \hat{m}_h(X_{t-1}))^2 \\ - \frac{\sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right) (X_t - \hat{m}_h(X_{t-1}))^2}{\sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right)}] = 0 \end{aligned}$$

zu (ii):

$$\begin{aligned} E^* \left[\frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x)) \right]^2 \\ = \frac{1}{T^2 h} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) ((X_t - \hat{m}_h(X_{t-1}))^2 - \hat{\sigma}_h^2(x))^2 \\ = \frac{1}{T^2 h} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) ((X_t - m(X_{t-1}) + m(X_{t-1}) - \hat{m}_h(X_{t-1}))^2 \\ - \sigma^2(X_{t-1}) + \sigma^2(X_{t-1}) - \hat{\sigma}_h^2(x))^2 \\ = \frac{1}{T^2 h} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}))^2 \end{aligned}$$

$$\begin{aligned}
& + 2(X_t - m(X_{t-1}))(m(X_{t-1}) - \hat{m}_h(X_{t-1})) + (m(X_{t-1}) - \hat{m}_h(X_{t-1}))^2 \\
& + \sigma^2(X_{t-1}) - \hat{\sigma}_h^2(x))^2
\end{aligned}$$

Wie man mit Hilfe von (iii) sehen wird, ist dieser Ausdruck nach Wahrscheinlichkeit beschränkt.

zu (iii):

$$\begin{aligned}
s_n^2 &= \sum_{t=1}^T E^* \left[\frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x)) \right]^2 \\
&= \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) ((X_t - m(X_{t-1}))^2 - \sigma^2(X_{t-1}) \\
&\quad + 2(X_t - m(X_{t-1}))(m(X_{t-1}) - \hat{m}_h(X_{t-1})) + (m(X_{t-1}) - \hat{m}_h(X_{t-1}))^2 \\
&\quad + \sigma^2(X_{t-1}) - \hat{\sigma}_h^2(x))^2
\end{aligned}$$

Von den nun nach Ausmultiplizieren auftretenden Termen wird exemplarisch nur der folgende untersucht. Die anderen können analog oder ähnlich wie in vorhergehenden Kapiteln betrachtet werden.

$$\begin{aligned}
& \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) (m(X_{t-1}) - \hat{m}_h(X_{t-1}))^2 (\sigma^2(X_{t-1}) - \hat{\sigma}_h^2(x)) \\
&= \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) m^2(X_{t-1}) \sigma^2(X_{t-1}) \\
&\quad - 2 \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) m(X_{t-1}) \hat{m}_h(X_{t-1}) \sigma^2(X_{t-1}) \\
&\quad + \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) \hat{m}_h^2(X_{t-1}) \sigma^2(X_{t-1}) \\
&\quad - \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) m^2(X_{t-1}) \hat{\sigma}_h^2(x) \\
&\quad + 2 \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) m(X_{t-1}) \hat{m}_h(X_{t-1}) \hat{\sigma}_h^2(x) \\
&\quad - \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) \hat{m}_h^2(X_{t-1}) \hat{\sigma}_h^2(x) \\
&\rightarrow 0
\end{aligned}$$

denn bekanntermaßen gilt $\hat{\sigma}_h^2(x) \rightarrow \sigma^2(x)$ nach Wahrscheinlichkeit, und für den zweiten Summanden (die anderen können analog betrachtet werden)

$$\begin{aligned}
& \frac{1}{Th} \sum_{t=1}^T K^2\left(\frac{x - X_{t-1}}{h}\right) m(X_{t-1}) \hat{m}_h(X_{t-1}) \sigma^2(X_{t-1}) \\
&= \frac{1}{h} \int K^2\left(\frac{x - u}{h}\right) m(u) \hat{m}_h(u) \sigma^2(u) du \\
&= \frac{1}{h} \int K^2(v) m(x + hv) \hat{m}_h(x + hv) \sigma^2(x + hv) h dv \\
&\rightarrow m^2(x) \sigma^2(x) \int K^2(v) dv
\end{aligned}$$

denn $\hat{m}_h(x) \rightarrow m(x)$ nach Wahrscheinlichkeit.

Setzt man die Ergebnisse aus dem Verhalten der anderen Terme zusammen erhält man die behauptete Konvergenz.

Insgesamt ergibt sich

$$s_n^2 \rightarrow \sigma^4(x) p(x) \int K^2(v) dv (\kappa(x) - 1) \text{ nach Wahrscheinlichkeit.}$$

Die Lindebergbedingung ist erfüllt:

$$\begin{aligned}
& \frac{1}{s_n^2} \sum_{t=1}^T \int 1 \left\{ \frac{1}{\sqrt{Th}} K\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x)) \geq \epsilon s_n \right\} \\
& \quad \frac{1}{Th} K^2\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^2 dP^* \\
& \leq \frac{1}{hs_n^2} \left[\int 1 \left\{ K^2\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^2 \geq \epsilon^2 s_n^2 Th \right\} dP^* \right. \\
& \quad \left. \int K^4\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^4 dP^* \right]^{\frac{1}{2}} \\
& = \frac{1}{hs_n^2} [P^* \{ K^2\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^2 \geq \epsilon^2 s_n^2 Th \} \\
& \quad E^* K^4\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^4]^{\frac{1}{2}} \\
& \leq \frac{1}{hs_n^2} \left(\frac{1}{\epsilon^4 s_n^4 T^2 h^2} \left(E^* K^4\left(\frac{x - X_{t-1}^*}{h}\right) ((X_t^* - \hat{m}_h(X_{t-1}^*))^2 - \hat{\sigma}_h^2(x))^4 \right)^2 \right)^{\frac{1}{2}}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{hs_n^2} \frac{1}{\epsilon^2 s_n^2 Th} h \frac{1}{Th} \sum_{t=1}^T K^4\left(\frac{x - X_1}{h}\right) ((X_2 - \hat{m}_h(X_1))^2 - \hat{\sigma}_h^2(x))^4 \\
&= O\left(\frac{1}{Th}\right) O(1) = o(1)
\end{aligned}$$

Da der Erwartungswert hier aber zu null geschätzt wird, ist eine explizite Erwartungswertschätzung mittels einer (Über-)Glättung des Schätzers $\hat{\sigma}_h^2(x)$ durch Wahl einer größeren Bandweite g und mit Hilfe der Kernfunktion K nötig. Dann ergibt sich der Schätzer:

$$\tilde{\sigma}_g^2(x) := \frac{\sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right) \hat{\sigma}_g^2(X_{t-1})}{\sum_{t=1}^T K\left(\frac{x - X_{t-1}}{h}\right)}$$

Satz 5.2: Unter den Voraussetzungen aus dem letzten Satz und (W5) gilt

$$\sqrt{Th}(\hat{\sigma}_h^{*2}(x) - \hat{\sigma}_h^2(x) + \tilde{\sigma}_g^2(x) - \hat{\sigma}_g^2(x)) \implies \mathcal{N}(b(x), \tau^2(x)) \quad \text{nach Wahrscheinlichkeit,}$$

wobei

$$\begin{aligned}
b(x) &= B \int v^2 K(v) dv \left[\frac{(\sigma^2)'(x) p'(x)}{p(x)} + \frac{1}{2} (\sigma^2)''(x) \right] \\
\tau^2(x) &= \frac{\sigma^4(x)}{p(x)} \int K^2(v) dv (\kappa(x) - 1)
\end{aligned}$$

Beweis: Aufgrund des vorigen Satzes bleibt zu zeigen: $\sqrt{Th}(\tilde{\sigma}_g^2(x) - \hat{\sigma}_g^2(x)) \rightarrow b(x)$ nach Wahrscheinlichkeit.

Einsetzen von $\tilde{\sigma}_g^2(x)$ und Taylorentwicklung von $\hat{\sigma}_g^2(X_{t-1})$ liefern:

$$\begin{aligned}
&\sqrt{Th}(\tilde{\sigma}_g^2(x) - \hat{\sigma}_g^2(x)) \\
&= \frac{\frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (X_{t-1} - x) (\hat{\sigma}_g^2)'(x)}{\frac{1}{Th} K\left(\frac{x - X_{t-1}}{h}\right)} \\
&\quad + \frac{\frac{1}{2} \frac{1}{\sqrt{Th}} \sum_t K\left(\frac{x - X_{t-1}}{h}\right) (X_{t-1} - x)^2 (\hat{\sigma}_g^2)''(\hat{X}_{t-1})}{\frac{1}{Th} K\left(\frac{x - X_{t-1}}{h}\right)}
\end{aligned}$$

Der letzte Ausdruck ist derselbe Term wie der Biasterm im Wild Bootstrap und kann genauso behandelt werden wie im Kapitel 2.1. Damit ist der Satz bewiesen. \square

Kapitel 6

Autoregressives Bootstrap

Eine dritte Möglichkeit, die Verteilung $\mathcal{L}^\sigma(x)$ konsistent zu schätzen, stammt aus u.a. Franke, Kreiß, Mammen (2001) und soll in diesem Kapitel kurz vorgestellt werden.

Wir betrachten wieder einen stationären und geometrisch ergodischen Prozess, das NARCH-Modell

$$X_t = m(X_{t-1}) + \sigma(X_{t-1})\epsilon_t, \quad t = 0, 1, 2, \dots \quad (6.1)$$

Folgende Voraussetzungen werden benötigt, um die Beweise einfach zu halten:

(A1) Die Zeitreihe X_1, \dots, X_T ist eine Realisierung der stationären Version des obigen NARCH-Prozesses mit unabhängig identisch verteilten Residuen $\epsilon_1, \dots, \epsilon_T$ mit Erwartungswert null und Varianz eins. Die Funktion m ist Lipschitz-stetig mit Lipschitz-Konstante L_m und die Funktion σ ist Lipschitz-stetig mit Lipschitz-Konstante L_σ .

(A2) Es gilt $L_m + L_\sigma E|\epsilon_1| < 1$.

(A3) Es gibt eine Konstante $\sigma_0 > 0$ derart, dass $\sigma(x) > \sigma_0$ für alle $x \in \mathbb{R}$.

(A4) Die Verteilung P_ϵ der Residuen ϵ_t besitzt eine Dichte p_ϵ mit folgenden Eigenschaften: $\inf_{x \in C} p_\epsilon(x) > 0$ für alle kompakten Mengen C , p_ϵ ist zweimal

stetig differenzierbar mit beschränkten Ableitungen und $\sup_{x \in \mathbb{R}} |xp'_\epsilon| < \infty$. Außerdem gilt für eine Konstante $\gamma > 2$ $E|\epsilon_1|^\gamma < \infty$.

(A5) m ist zweimal stetig in x differenzierbar mit beschränkten Ableitungen.

(A6) K besitzt kompakten Träger $[-1, 1]$. K ist symmetrisch, besitzt eine beschränkte Ableitung auf $(-1, 1)$ und $\int K(v)dv = 1$. Für die Bandweite h gilt $h \rightarrow 0$ und $Th^5 \rightarrow B^2 > 0$.

(A7) Die Bootstrap-Residuen ϵ_t^* besitzen (bedingten) Erwartungswert null und Varianz eins und es gilt $d_K(P_{\epsilon}, P_\epsilon) = o_P(1)$. Für den Startwert X_0^* des Bootstrap-Prozesses gilt $E^*[X_0^*] = O_P(1)$. $E^*|\epsilon_t^*|^\gamma = O_P(1)$ für eine Konstante $\gamma > \frac{5}{2}$.

(A8) Es existiert eine Folge $\gamma_T \rightarrow \infty$ derart, dass $\sup_{|x| > \gamma_T} |\tilde{m}(x)| = O_P(1)$, $\sup_{|x| > \gamma_T} |\tilde{\sigma}(x)| = O_P(1)$ und $\sup_{|x| > \gamma_T} |\tilde{\sigma}^{-1}(x)| = O_P(1)$.

(A9) Mit γ_T aus (A8) gilt $\sup_{x \leq \gamma_T} |\tilde{\sigma}(x) - \sigma(x)| = o_P(\gamma_T^{-1})$. Für alle $C > 0$ und für $j = 0$ und $j = 1$ gilt $\sup_{x \leq \gamma_T} |\tilde{m}^{(j)}(x) - m^{(j)}(x)| = o_P(1)$ und $\sup_{x \leq C\gamma_T} |\tilde{p}_\epsilon^{(j)}(x) - p_\epsilon^{(j)}(x)| = o_P(1)$.

(A10) $\sup_{|x-x_0| \leq h} |\tilde{m}^{(2)}(x) - m^{(2)}(x)| = o_P(1)$.

Für das autoregressive Bootstrap des Varianz-Schätzers benötigen wir die folgenden zusätzlichen Annahmen:

(A11) Es gilt $E|\epsilon_1|^\gamma < \infty$ für eine Konstante $\gamma > 4$. Das bedingte vierte Moment $E[\epsilon_1^4 | X_1 = x]$ ist eine stetige Funktion in $x = x_0$. Die Funktion σ ist zweimal stetig in x differenzierbar mit beschränkten Ableitungen. Für die Bandweite h' gilt $h' \rightarrow 0$ und $T(h')^5 \rightarrow (B')^2 > 0$.

(A12) $E^*|\epsilon_t^*|^\gamma = O_P(1)$ für eine Konstante $\gamma > 5$. $E^*|\epsilon_t^*|^4 = E|\epsilon_t|^4 + o_P(1)$. Für $j = 0$ und $j = 1$ gilt $\sup_{x \leq \gamma_T} |\tilde{\sigma}^{(j)}(x) - \sigma^{(j)}(x)| = o_P(\gamma_T^{-1})$. $\sup_{|x-x_0| \leq h} |\tilde{\sigma}^2(x) - \sigma^2(x)| = o_P(1)$.

Wie schon in der Einleitung erwähnt wird beim autoregressiven Bootstrap das NARCH-Modell exakt nachgebildet:

$$X_t^* = \tilde{m}(X_{t-1}^*) + \tilde{\sigma}(X_{t-1}^*)\epsilon_t^*,$$

wobei \tilde{m} und $\tilde{\sigma}$ Schätzer von m und σ sind und $\epsilon_1^*, \dots, \epsilon_T^*$ eine (unter der ursprünglichen Stichprobe X_0, \dots, X_T) bedingt unabhängig identisch verteilte Stichprobenwiederholung (Resample).

Im Folgenden wird die Wahl der Schätzer \tilde{m} , $\tilde{\sigma}^2$ und \tilde{P}_ϵ erörtert. Für die Schätzer \tilde{m} und $\tilde{\sigma}^2$ können trunkierte Nadaraya-Watson-Schätzer gewählt werden und für die Verteilung \tilde{P}_ϵ die empirische Verteilung der Residuen. Für Konstanten C_m , C_σ und c_σ setzen wir $\bar{C}_m = C_m|\bar{X}|$, $\bar{C}_\sigma = C_\sigma\bar{r}^2$ und $\bar{c}_\sigma = c_\sigma\bar{r}^2$, wobei \bar{X} der Stichprobenmittelwert und \bar{r}^2 das arithmetische Mittel der quadrierten Residuen $\hat{r}_t^2 = (X_t - \hat{m}_h(X_{t-1}))^2$ ist. Wir setzen

$$\tilde{m}(x) = \begin{cases} \hat{m}_g(x) & \text{für } |\hat{m}_g(x)| \leq \bar{C}_m \\ \bar{C}_m & \text{für } \hat{m}_g(x) > \bar{C}_m \\ -\bar{C}_m & \text{für } \hat{m}_g(x) < -\bar{C}_m \end{cases}$$

$$\tilde{\sigma}(x) = \begin{cases} \hat{\sigma}_{j,g'}(x) & \text{für } |\hat{\sigma}_{j,g'}(x)| \leq \bar{C}_\sigma \\ \bar{C}_\sigma & \text{für } \hat{\sigma}_{j,g'}(x) > \bar{C}_\sigma \\ \bar{c}_\sigma & \text{für } \hat{\sigma}_{j,g'}(x) < -\bar{c}_\sigma \end{cases},$$

wobei $j = 1$ oder 2 und wobei g und g' von h und h' verschiedene Bandweiten sind, z.B. wird g im Gegensatz zu h so gewählt, dass $Tg^5 \rightarrow \infty$. Das bedeutet eine Überglättung des Schätzers wie auch beim Wild Bootstrap. Wir definieren nun \tilde{P}_ϵ . Für eine Folge $\gamma_T \rightarrow \infty$ setzen wir $I_T = \{1 \leq t \leq T : |\hat{r}_t| \leq \gamma_T\}$. Dann wird \tilde{P}_ϵ definiert als empirische Verteilung von $\hat{r}_t - \bar{r}_{I_T}$ für $t \in I_T$. Hierbei ist \bar{r}_{I_T} das arithmetische Mittel der \hat{r}_t für $t \in I_T$.

Der Bootstrap-Schätzer von σ wird nun folgendermaßen definiert:

$$\hat{\sigma}_{h'}^{*2}(x) = \frac{\frac{1}{Th'} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h'}\right) (X_t^* - \hat{m}_{h'}(X_{t-1}^*))^2}{\frac{1}{Th'} \sum_{t=1}^T K\left(\frac{x - X_{t-1}^*}{h'}\right)}.$$

Die bedingte Verteilung von $\sqrt{Th}\{\hat{\sigma}_h^{*2}(x_0) - \tilde{\sigma}^2(x_0)\}$, gegeben X_0, \dots, X_T , wird mit $\mathcal{L}_B^\sigma(x)$ bezeichnet. Dieses ist der Bootstrap-Schätzer von $\mathcal{L}^\sigma(x)$, der Verteilung von $\sqrt{Th}\{\hat{\sigma}_h^2(x) - \sigma^2(x)\}$.

Wir kommen nun zum Hauptergebnis dieses Kapitels, der Konsistenz des Bootstrap-Schätzers $\mathcal{L}_B^\sigma(x)$.

Satz 6.1: Gelten (AB1) - (AB12). Dann gilt für alle $x_0 \in \mathbb{R}$:

$$d_K(\mathcal{L}_B^\sigma(x_0), \mathcal{L}^\sigma(x_0)) \rightarrow 0 \text{ nach Wahrscheinlichkeit}$$

d_K bezeichnet den Kolmogorov-Abstand, d.h. für zwei Verteilungen P und Q ist der Abstand $d_K(P, Q)$ definiert als $\sup_{x \in \mathbb{R}} |P(X \leq x) - Q(X \leq x)|$

Beweis: siehe Franke, Kreiß, Mammen (2001).

Kapitel 7

Anhang

7.1 Mischen

Wir stellen eine Definition für starkes Mischen und einige in den vorhergehenden Kapiteln verwendete Eigenschaften von stark mischenden Zufallsfolgen vor. Für Einzelheiten verweisen wir auf Doukhan (1994).

Sei $\{X_t : t = 0, 1, 2, \dots\}$ eine Folge von \mathbb{R}^d -wertigen Zufallsvariablen, $d = 1, 2, \dots$, und \mathcal{F}_n^m die von $\{X_t : n \leq t \leq m\}$, $0 \leq n \leq m$ erzeugte σ -Algebra.

$\{X_t\}_{t \geq 0}$ heisst *stark mischend* oder α -*mischend*, falls für die Mischkoeffizienten

$$\alpha_k = \sup_n \sup_{A \in \mathcal{F}_0^n, B \in \mathcal{F}_{n+k}^\infty} |P(A \cap B) - P(A)P(B)|$$

gilt

$$\lim_{k \rightarrow \infty} \alpha_k = 0.$$

Falls $\{X_t\}$ stark mischend ist, dann auch $\{Y_t\}$ mit $Y_t = f_t(X_t)$ mit denselben oder kleineren Mischkoeffizienten (siehe Doukhan (1994), S.15).

Falls ein Markov-Prozess $\{X_t\}_{t \geq 0}$ stationär und geometrisch ergodisch ist, dann ist $\{X_t\}$ geometrisch stark mischend mit geometrisch fallenden Mischkoeffizienten $\alpha_k \leq C\rho^k$, $C > 0$, $0 < \rho < 1$ (siehe Doukhan (1994), S.89).

Aus Doukhan (1994), Theorem 3, S.9 haben wir die folgende Kovarianzgleichung:

Lemma: Seien X und Y bezüglich der σ -Algebren \mathcal{U} und \mathcal{V} messbare Zufallsvariablen. Dann gilt für $p, q, r \geq 1$ mit $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$

$$|Cov(X, Y)| \leq 8\alpha^{\frac{1}{r}}(\mathcal{U}, \mathcal{V}) \|X\|_p \|Y\|_q,$$

wobei $\alpha(\mathcal{U}, \mathcal{V}) = \sup \{|P(U)P(V) - P(U \cap V)| : U \in \mathcal{U}, V \in \mathcal{V}\}$.

7.2 Martingaldifferenzschemata

Wir geben eine Definition für Martingaldifferenzschemata und präsentieren einen zentralen Grenzwertsatz von Brown (1971) und eine Ungleichung für Martingaldifferenzschemata aus z.B. Hall und Heyde (1980).

Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathcal{F}_n : n \in \mathbb{N}_0)$ eine Folge von σ -Algebren mit $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{A}$. Sei $(\xi_n : n \in \mathbb{N}_0)$ eine Folge von Zufallsvariablen auf (Ω, \mathcal{A}, P) , wobei ξ_n messbar bezüglich \mathcal{F}_n ist. Die stochastische Folge $[(\xi_n, \mathcal{F}_n) : n \in \mathbb{N}_0]$ heißt *Martingaldifferenzschema*, falls

$$E|\xi_n| < \infty \quad \text{und} \quad E[\xi_{n+1}|\mathcal{F}_n] = 0 \quad P\text{-f.s.} \quad \text{für alle } n \in \mathbb{N}_0.$$

Burkholder Ungleichung: Falls $\{X_t : 1 \leq t \leq T\}$ ein Martingaldifferenzschema ist und $1 < p < \infty$, dann existieren von p abhängige Zahlen C_1 und C_2 derart, dass

$$C_1 E \left| \sum_{t=1}^T X_t^2 \right|^{\frac{p}{2}} \leq E \left| \sum_{t=1}^T X_t \right|^p \leq C_2 E \left| \sum_{t=1}^T X_t^2 \right|^{\frac{p}{2}}$$

Zentraler Grenzwertsatz für Martingaldifferenzschemata: Sei $[(\xi_n, \mathcal{F}_n) : n \in \mathbb{N}_0]$ ein Martingaldifferenzschema mit $E\xi_n^2 < \infty$ für alle $n \in \mathbb{N}_0$. Sei $\mathcal{X} := [(X_n, \mathcal{F}_n) : n \in \mathbb{N}_0]$ mit $X_n := \frac{1}{\sqrt{n}} \sum_{k=1}^n \xi_k$, wobei $X_0 = 0$ und $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Falls die bedingte Lindeberg-Bedingung

$$\frac{1}{n} \sum_{k=1}^n E[\xi_k^2 \mathbf{1}_{\{|\xi_k| > \epsilon \sqrt{n}\}} | \mathcal{F}_{k-1}] \rightarrow 0 \quad \text{nach Wahrscheinlichkeit für alle } \epsilon > 0$$

$$\text{erfüllt ist und falls} \quad \frac{1}{n} \sum_{k=1}^n E[\xi_k^2 | \mathcal{F}_{k-1}] \rightarrow \sigma^2 \quad \text{nach Wahrscheinlichkeit,}$$

dann gilt $X_n \Rightarrow \mathcal{N}(0, \sigma^2)$.

Literaturverzeichnis

- [1] Bachelier, Louis: Théorie de la spéculation, *Annales de l'École Normale Supérieure*, **III-17**, 21-86 (1900)
- [2] Bernstein, Sergey N.: Sur l'Extension du théorème du calcul des probabilités aux Sommes de Quantités Dependantes, *Mathematische Annalen* **97**, 1-59 (1927)
- [3] Black, Fischer; Scholes, Myron: The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* **81**, 637-654 (1973)
- [4] Bollerslev, Tim: Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, **31**, 307-327 (1986)
- [5] Bosq, Denis: *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statistics **110**, Springer, New York (1996)
- [6] Brown, Bruce M.: Martingale Central Limit Theorems, *Annals of Mathematical Statistics* **42**, 54-66 (1971)
- [7] Diebolt, J.; Guegan, D.: Probabilistic properties of the General Nonlinear Markovian Process of Order One and Applications to Time Series Modelling. Rapport Technique de L.S.T.A. 125, Université Paris XIII
- [8] Doukhan, Paul: *Mixing: Properties and Examples*, Lecture Notes in Statistics **85**, Springer, New York (1994)
- [9] Drees, Holger; Stărică, Cătălin: A Simple Non-stationary Model for Stock Returns, *Preprint* (2002)
- [10] Efron, Bradley: Bootstrap Methods: Another Look at the Jackknife, *Annals of Statistics* **7**, No.1, 1-26 (1979)

- [11] Einstein, Albert: On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-kinetic Theory of Heat, *Annals of Physics* **17** (1905)
- [12] Engle, Robert F.: Autoregressive Conditional Heteroskedastic Models with estimates of the Variance of United Kingdom Inflation, *Econometrica* **50**, 987-1007 (1982)
- [13] Fama, Eugene F.: The Behavior of Stock Market Prices, *Journal of Business* **38**, 34-105 (1965)
- [14] Fan, Jianqing; Gijbels, Irène: *Local Polynomial Modelling and Its Applications*, Chapman and Hall, New York (1995)
- [15] Fan, Jianqing; Yao, Q.: Efficient Estimation of Conditional Variance Functions in Stochastic Regression, *Biometrika* **85**, 645-660 (1998)
- [16] Feller, William: *An Introduction to Probability Theory and its Applications, Vol. I*, Wiley, New York (1968)
- [17] Franke, Jürgen; Kreiß, Jens-Peter; Mammen, Enno: Bootstrap of Kernel Smoothing in Nonlinear Time Series, Discussion Paper SFB Quantifikation und Simulation ökonomischer Prozesse (1997)
- [18] Franke, Jürgen; Kreiß, Jens-Peter; Mammen, Enno: Bootstrap of Kernel Smoothing in Nonlinear Time Series, *Bernoulli* **8**, 1-37, (2002)
- [19] Franke, Jürgen; Wendel, M.: A Bootstrap Approach for Nonlinear Autoregressions - Some Preliminary Results. Aus: Jäkel, K.-H.; Rothe, G.; Sendler, W. eds.: Bootstrapping and Related Techniques. Lecture Notes in Economics and Mathematical Systems **376**, Springer, New York (1992)
- [20] Glosten, L., Jagannathan, R. and Runkle, D.: Relationship between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, **48**, 1779-1801 (1993)
- [21] Gouriéroux, C. and Monfort, A.: Qualitative threshold ARCH models, *Journal of Econometrics*, **52**, 159-199 (1992)
- [22] Györfy, L.; Härdle, Wolfgang; Sarda, P.: *Nonparametric Curve Estimation from Time Series*, Lecture Notes in Statistics **60**, Springer, New York (1989)

- [23] Härdle, Wolfgang: *Smoothing Techniques: With Implementation in S*, Springer, New York (1991)
- [24] Härdle, Wolfgang; Mammen, Enno: Comparing Nonparametric versus Parametric Regression Fits, *Annals of Statistics* **21**, 1926-1947 (1994)
- [25] Hall, Peter; Heyde, C.: *Martingale Limit Theory and Its Applications*, Academic Press, New York (1980)
- [26] Herzel, Stefano; Stărică, Cătălin; Tütüncü, Reha: A Non-stationary Multivariate Model for Financial Returns, *Preprint* (2003)
- [27] Ibragimov, I.: Some Limit Theorems for Stationary Processes, *Theory of Probability and Its Applications*, **7**, 349-382 (1962)
- [28] Kreiß, Jens-Peter: Nonparametric Estimation and Bootstrap for Financial Time Series. Aus: Chan, W.S.; Li, W.K.; Tong, H. (Hrsg.): *Statistics and Finance: An Interface*, Imperial College Press, London (2000)
- [29] Kreiß, Jens-Peter; Neumann, Michael H.; Yao, Q.: Bootstrap Tests for Simple Structures in Nonparametric Time Series Regression, *Preprint* (2002)
- [30] Kreutzberger, Eva: *Bootstrap für nichtlineare AR(1)-Prozesse*, Dissertation, Universität Kaiserslautern (1993)
- [31] Mammen, Enno: *When Does Bootstrap Work ?* Lecture Notes in Statistics **77**, Springer, New York (1992)
- [32] Mandelbrot, Benoit: The Variation of Certain Speculative Prices, *Journal of Business*, **36**, 394-419 (1963)
- [33] Masry, E.; Tjøstheim, D.: Nonparametric Estimation and Identification of Nonlinear ARCH Time Series: Strong Convergence and Normality, *Econometric Theory*, **11**, 258-289 (1994)
- [34] Merton, Robert C.: The Theory of Rational Option Pricing, *Bell Journal of Economic Management Sciences*, **4**, 141-183 (1973)
- [35] Meyn, S.P.; Tweedie, R.L.: *Markov Chains and Stochastic Stability*, Springer, New York (1993)
- [36] Mikosch, Thomas: Modeling Dependence and Tails of Financial Time Series. Aus: Finkenstädt, B.; Rootzén, B. (Hrsg.) *Extreme Values in Finance, Telecommunications and the Environment*, Chapman and Hall, 185-286 (2003)

- [37] Mikosch, Thomas; Stărică, Cătălin: Long Range Dependence Effects and ARCH Modeling. Aus: Doukhan, P.; Oppenheim G.; Taqqu, M.S. (Hrsg.) *Long Range Dependence*. Birkhäuser, Boston (2003)
- [38] Neumann, Michael H.; Kreiß, Jens-Peter: Regression-type Inference in Nonparametric Autoregression, *Annals of Statistics* **26**, 1570-1613 (1998)
- [39] Polzehl, Jörg; Spokoiny, Vladimir; Stărică, Cătălin: When Did the 2001 Recession Really Start ?, *Preprint* (2004)
- [40] Priestley, M.B.: *Nonlinear and Nonstationary Time Series Analysis*, Academic Press, London (1988)
- [41] Politis, Dimitris N.; Romano, Joseph P.; Wolf, Michael: *Subsampling*, Springer, New York (1999)
- [42] Robinson, Peter M.: Nonparametric Estimators of Time Series, *Journal of Time Series Analysis*, **4**, 185-207 (1983)
- [43] Rosenblatt, M.: A Central Limit Theorem and a Strong Mixing Condition, *Proceedings National Academic Sciences USA*, **42**, 43-47 (1956)
- [44] Samuelson, Paul A.: Mathematics of Speculative Prices, *SIAM Review*, **15**, 1-42 (1973)
- [45] Straumann, Daniel: *Estimation in Conditionally Heteroskedastic Time Series Models*, PhD Thesis, University of Copenhagen (2003)
- [46] Taylor, Stephen: *Modelling Financial Time Series*, Wiley, Chichester (1986)
- [47] Tong, H.: *Nonlinear Time Series. A Dynamic Approach*, Oxford University Press, Oxford (1990)
- [48] Wu, C.F.J.: Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis, *Annals of Statistics* **14**, 1261-1343 (1986)
- [49] Zakoian, J. M.: Threshold heteroskedastic models, *Journal of Economic Dynamic Control*, **18**, 931-955 (1994)

Zusammenfassung

In der vorliegenden Arbeit werden bedingt heteroskedastische Zeitreihen der Ordnung eins betrachtet.

Das zweite Kapitel beschäftigt sich mit Kurvenschätzung in nichtlinearen autoregressiven bedingt heteroskedastischen Prozessen der Ordnung eins. Die bedingte Varianz wird nichtparametrisch durch Nadaraya-Watson-Schätzer approximiert. Asymptotische Normalität wird im nichtparametrischen Modell nachgewiesen. Die Konsistenz der ersten und zweiten Ableitung des Nadaraya-Watson-Schätzers für die bedingte Varianz wird gezeigt.

In folgenden Kapiteln betrachten wir Bootstrap-Methoden in nichtlinearen autoregressiven bedingt heteroskedastischen Modellen der Ordnung eins. Für den Nadaraya-Watson-Schätzer der bedingten Varianz werden Bootstrap-Schätzer entwickelt. Die behandelten Bootstrap-Methoden sind das sogenannte Wild, das paarweise und das autoregressive Bootstrap. Wir zeigen die Konsistenz der Verfahren. Für das Wild und das paarweise Bootstrap wird desweiteren gezeigt, dass diese Verfahren auch ohne Annahme eines zugrundeliegenden nichtlinearen autoregressiven bedingt heteroskedastischen (NARCH(1)) Prozesses konsistent sind.